Macroeconomic conditions, corporate default, and default clustering

Article

Accepted Version

Xing, Kai, Luo, Dan and Liu, Lanlan (2023) Macroeconomic conditions, corporate default, and default clustering. Economic Modelling, 118. 106079. ISSN 0264-9993 doi: https://doi.org/10.1016/j.econmod.2022.106079 Available at https://centaur.reading.ac.uk/109180/

To link to this article DOI: http://dx.doi.org/10.1016/j.econmod.2022.106079

Publisher: Elsevier

Macroeconomic Conditions, Corporate Default, and Default Clustering

Kai Xing[a], Dan Luo[b], Lanlan Liu[c][1]

[a] *School of Economics and Management, Nanchang University, Nanchang 330031, China*

[b] *Henley Business School, University of Reading, Reading RG6 6UD, UK*

[c] *Business School, Soochow University, Suzhou 215012, China*

Abstract

This study thoroughly investigates how, and to what extent, macroeconomic conditions interact with corporate default in the US industrial sector from 1980 to 2014. Using an extensive data set of macro-level and micro-level variables, we construct five categories of indicators and measure macroeconomic conditions by investigating co-movements within each category of indicator. We find macroeconomic conditions have bidirectional causal interaction with corporate defaults across different economic regimes, reflecting the existence of feedback causality. Moreover, we find that macroeconomic indicator constructed using the least absolute shrinkage and selection operator (LASSO) approach shows superior explanatory power as well as predictive power for default clustering, indicating that movements of these indicators cause correlated changes in firms' default rates. Overall, our study provides support for literature on default probability estimation from a macroeconomic perspective.

JEL classification: G01, G10, G33, E32.

Keywords: Macroeconomic conditions; macro indicator; corporate default; default clustering; default prediction

We are grateful to the Co-Editor, Sushanta Mallick, the guest editor, and two anonymous referees for their insightful comments and suggestions. The usual disclaimer applies.

[*] Corresponding author: Email: lanlanliu@suda.edu.cn

1. Introduction

The US economy has experienced a few severe clusters of corporate default over the past four decades, such as the savings and loan crisis in early 1990s, the dotcom bubble in 2001, and the subprime crisis over 2007-09. Such default clustering typically closely follows the business cycle, which is driven by firms' joint exposure to macroeconomic factors that are correlated over time (Duffie, 2009). Thus, detecting structural change within macroeconomic system is crucial for understanding the sources of default clustering.

In this paper, we investigate the dynamic causal relationship between macroeconomic conditions and corporate default. We examine the potential to predict default clustering by investigating the evolution of corporate defaults in the US industrial firms and macroeconomic data generating process over the thirty-five-year period from January 1980 to December 2014. For this, we employ a cross-disciplinary theory from the Natural Sciences that builds generic indicators to explore critical transitions in a complex system, and to capture domino effects to measure dynamics of macroeconomic conditions (Scheffer et al., 2012; Battiston et al., 2016; Diks et al., 2019; Xing and Yang, 2019).

This study is mainly motivated by studies from threefold. First, the prediction of corporate default risk has been extensively investigated by many studies since the proposal of the first bankruptcy model by Altman (1968), but they all have a common weakness of applying theories for well-behaved systems that are not well behaved (Sornette and Woodard, 2010). Specifically, these models often assume that financial (credit) market is stable and past events can provide signal for the unforeseeable events in the future. Bernanke (1993) argues that the credit market is dynamic rather than stable, and the predictive power of various predictors should reflect the market's perception of default risk. It is suggested that when the market's perception represented by the economic environment to default risk changes, the predictor should change as well. As suggested by Bose et al., (2021), credit market development such as law enforcement can reduce financial constraints and probabilities of default. The individual best predictor for default clustering can be time-varying (Hwang, 2019). For some factors, they are only useful in certain set of conditions, while others may be useful in a different set. In other word, a fixed group of factors may not be appropriate for the prediction of corporate default consistently. This phenomenon can be explained by an recently increasing consideration of Adaptive Market Hypothesis (AMH) proposed by Lo (2004)[2], who declares that market efficiency varies over time and both the Efficient Market Hypothesis (EMH) and market inefficiencies co-exist in an intellectually consistent manner. That is, the predictability of any models will be time-varying. As a result, the construction of an efficient indicator which can adapt to macroeconomic fluctuations would be of critical importance in predicting default risk reliability. Recent studies suggest that agents in the complex economic system follow an adaptive learning process with correlated trend-following behaviors (Hommes, 2021).

Additionally, this study is motived by studies that examine the role of macroeconomic factors in predicting corporate defaults. Despite of the theoretical debates about the causality relationship between dynamics of macroeconomic conditions and corporate default, relatively little research has been done to test empirically the relationship of the two. Theoretically speaking, corporate default would impact negatively on the economy as it may lead to the loss of market confidence and the worsening of the overall macroeconomic environment. Gertler (1988) points out that credit market conditions play a central role in the propagation of cyclical fluctuations. A sharp increase in insolvencies and bankruptcies, rising real debt burdens, collapsing asset prices, and bank failures would generate significant impact on

[2] In order to reconcile the efficient Market Hypothesis (EMH) and behaviour finance, Lo (2004) proposes AMH based on equity market. Lo (2004) claims that based on evolutionary principals, the degree of market efficiency results from the biological perspective, such as the number of competitors in the market, the magnitude of profit opportunities available, and the adaptability of the market participants. In fact, this is also the same case in the credit market. Bernanke (1993) argues that the credit market is filled with imperfect and asymmetric information, which makes market inefficiency.

the economy, slowing down the economic activities in general (Bernanke et al., 1999). While on the other hand, a deteriorated macroeconomic condition may also impose downward pressure on corporate default as it may influence the credit risks and firms' financing decisions directly (Hackbarth et al., 2006). Inversely, clusters of corporate default may have disastrous economic effects such as a sudden drop of consumption (Gouriéroux et al., 2021). Therefore, in this paper, a thorough investigation of the causal relationship between the macroeconomic factors and the corporate default probabilities would be conducted.

More relevantly, this study builds on literature the complexity theory, that is to capture critical transitions in the complex system. According to Scheffer et al., (2012), critical transitions stand for sharp regime shifts that punctuate the usual fluctuations around trends in a complex system and they are often triggered by the unpredictable external shocks. They further point out that in a complex system, the occurrence of critical transitions implies the existence of interactions among system components and such interactions may lead to cascade or domino effects. Specifically, if the interactions among system components become stronger, the behaviour of these components may vary greatly or influence the functionality and/or operations of other components. As a result, a measurement capturing total interactions of the system for the prediction of critical transitions, and hence, the future change of the system is needed. Battiston et al. (2016) suggest that critical transitions do exist in economic systems, leading to system collapse and the subsequent occurrence of macroeconomic crises. Diks et al., (2019) test the predictability of financial market crash and verify that financial time series behave consistently with the complexity theory.

We aim to provide a mechanism to verify the recent theoretical findings that dynamic of macroeconomic conditions may impact default clustering. In 2001, the Conference Board published the book *Business Cycle Indicators Handbook* which argues that the US economy is continuously evolving and is far too complex to be summarized by one economic series. Specifically, Chen (2010) constructs a dynamic capital structure model to provide a rational mechanism for "credit contagion" and market timing of debt issuance. It is concluded that "these clustered defaults are due to a sudden deterioration of macroeconomic conditions that causes firms' default boundaries to jump up, so that those firms with cash flows below the new default boundaries will default simultaneously."[3] He also uses the timing of debt issuing to explain this clustering phenomenon. Rates of debt issuance are procyclical with the business cycle, which reaches to the peak level when the economy switches from a bad state to a good state as a large number of firms tend to issue debt simultaneously. Such clustering phenomenon can be interpreted as firms are exposing to the common or correlated risk factors such as interest rates, stock returns, GDP growth (Chava and Jarrow, 2004; Das et al., 2007; Duffie et al., 2007; Azizpour et al., 2018). As a result, the comovement among these macro factors can lead to associated changes in the conditional default rates of the firms while a strong economic growth would decrease the likelihood of default across the board. Meanwhile, several studies also identify that there is a time lag between economic recession and default clustering and high defaults seemingly follow the occurrence of economic recession (Koopman and Lucas, 2005; Koopman et al., 2009). Therefore, it seems that clustered defaults are highly correlated with macroeconomic conditions and this study would like to investigate this further.

Based on the theory of capturing critical transitions in the complex system, we initially construct macro indicators to represent macroeconomic conditions by measuring interactions among macroeconomic factors. We use Johansen's cointegration test to confirm that macro indicators cointegrate with corporate default probabilities, implying that macroeconomic conditions tend to have a long run-relationship with corporate default in US industrial films during the sample period.

[3] According to Leland (2004), A "default boundary" can be defined as a level of asset value which may vary with the shit of time. If the value of a firm' asset is below to this level, then this firm will default on its debt.

Next, we apply causality techniques to figure out that macro indicators either result in or from corporate default probabilities, suggesting that there is a dynamic interaction between macroeconomic conditions and corporate default over time. We then investigate whether macro indicators can predict default clustering, then investigate how much power macro indicators and NBER recession indicator have in explaining default clustering. We measure default clustering by removing the number of months with the less occurrence of corporate defaults. Here, we use several filtering methods to identify small defaults in the sample and finalize default clustering. We find that macro indicators and NBER recession indicator are highly significant in explaining and predicting default clustering even after lagging by 3 months.

In addition, we find that there is an interactive effect between each indicator and recession without lagging by 3 months, suggesting that the simultaneous influence of each indicator and recession on default clustering is additive. After lagging these variables by 3 months, they do not have any interactive effect on default clustering. This phenomenon may be interpreted as follows. Since there is a lagged effect between macroeconomy and default cycle (Couderc et al., 2008), then the simultaneous influence of each indicator and recession loses predictive ability on corporate defaults. These results show that macroeconomic conditions can predict default clustering, indicating that an early-warning system to predict default clustering based on macro indicators.

Our study contributes to the literature on the importance of detecting critical transitions in the macroeconomic system, which can better reflect the dynamic change of macroeconomic conditions. In practice, the validity of such measurement has been approved and found to be efficient in various nature sciences literature. For the macroeconomic system, it could also be regarded as a complex system with dynamic interactions existing among its components. Therefore, in this research, we would like to extend this known approach to capture the critical transitions in the system using a series of macroeconomic indicators. Meanwhile, our study adds to the literature on predicting default clustering from macroeconomic perspective, as we verify the predictive ability of our effective indicator. This can help with building an early-warning system for clustered default events in practice.

The remainder of the paper is organized as follows. Section 2 describes the data sets used and explains how they are constructed. Section 3 provides the combinations of macro factors which are used to build macro indicators for describing macroeconomic conditions. Section 4 details how to construct the indicators. Sections 5 shows the description of procedure of doing statistical analysis between corporate default and macroeconomic conditions. Section 5 exhibits empirical results, which is followed by Section 7 discussions. The final part is the conclusion.

2. The Data

This section is to present the data in this study, which includes measure of corporate default, measure of macroeconomic conditions, and summary of statistics, respectively.

2.1 Measure of corporate default

We focus on the default events of the listed and non-listed firms from industrial sectors. First, the reason why we use listed and non-listed firms together. Although default crises have been the focus of many studies, prior studies mainly focus on default crises in the listed firms. The main reason for this may simply be that the firm-specific information in the non-listed firms is not available to researchers. For example, Merton model and its extensions are only limited to listed firms since these models are based on equity-price information. However, Jacobson et al. (2013) argue that the privately held firms are typically responsible for over half of GDP in developed economies. Then, both listed firms and non-listed firms should be considered simultaneously, which can reflect the big picture of corporate default in the industrial economic sector in US. Second, the primary reason for selecting industrial firms is that other economic sectors such as media, financial sectors, gas and electric utilities are strongly influenced by regulators, affecting the return on equity, revenues and thus the risk of the default (Eom et al., 2004).

Corporate default information for US industrial firms is extracted from Moody's Default Recovery Database (DRD). DRD provides default histories for all rated US industrial firms, including a list of monthly defaulters and of withdrawn ratings. This information can be used to construct monthly issuer-weighted default probabilities (hereafter referred to as IDPs) for measuring the clustering of default risk at the market level. IDPs are defined as the number of monthly defaults over the rated firms in the beginning of each month, after removing the number of credits in each month.

Following Duffie et al. (2009), we further clean the default data in order to remove dependent default events. We first consider the influence of family firms may default together with their subsidiaries. DRD contains multiple entries for these firms. Then a parent's consolidated financial information is used to study the default decision for the whole family of firms. If a firm consecutively defaults more than once within two years, these multiple default events are counted as a single default event and the first default event is regarded as the default date for the firm.

Since the quality of default records was not reliable from 1898 to 1980 (Lando and Nielsen, 2010), we only use data from January 1980 to December 2014. After data cleaning, this leaves us with a total of 10368 firms covering listed firms and privately held firms, which comprise 1842 defaults.[4] We use this information to calculate the timeseries of IDPs for US industrial firms. Figure 1 shows the time series of total defaults, total exposures, and IDPs for US industrial firms. We find that default clustering occurs around the recession years of 1990, 2001, and the financial crisis of 2007-2009 in the graph of total defaults.

2.2 Measure of Macroeconomic conditions

We collect data of 114 macro (macroeconomic) factors from Federal Reserve Economic Database (FRED). Description of data is presented in Table 1. We select series following Koopman et al. (2011) to depict macroeconomic conditions for default risk. These series are classified into 2 categories. One reflects macro-level conditions, the other micro-level conditions. 114 macro factors are monthly time series from January 1980 to December 2014. 89 macro factors are all directly downloaded from FRED without any transformation from the other frequency to the monthly type. The rest are quarterly and annual, such as gross domestic product (GDP), gross national product (GNP), and Federal debt: total public debt etc. Cubic spline interpolation is used for converting quarterly and annual data to monthly data (Jordan and Jordan, 1997). All variables except exchange rates, interest rates, and industrial material price index are seasonally adjusted, and they are all transformed to growth rate except four yield spread ratios namely T10YFF, SPREAD.GS, SPREAD.MOODY.1, and SPREAD.MOODY.2.

We follow Scheffer et al. (2012) and Xing and Yang (2019) to construct macro indicators so as to measure macroeconomic conditions. The concept of constructing macro indicator can be explained as the interaction between nodes in a network or rising temporal interaction. The nodes in the macroeconomic system represent macroeconomic variables. The network here represents the interaction among these variables.

Basically, there are two major issues on constructing macro indicator for measuring macroeconomic conditions. On the one hand, macro indicator is calculated by macroeconomic system components, then how to better understand the system components improves predicting performance of macro indicator on capturing macroeconomic conditions; on the other hand, the method of measuring interaction among system components also plays an important role on improving the performance of macro indicator. Next, we briefly show the variable selection of macroeconomic system, suggesting how to construct the

[4] According to the DRD, Moody's started to record defaults in 1898.

network in the macroeconomic system, which is followed by the description of how to measure the interactions among macroeconomic system components.

2.2.1 Variable selection approaches

The abovementioned macroeconomic variables represent system components, which are classified into five types of macroeconomic system. These are shown in Table 2. Descriptions are given next.

(1) Augmented macroeconomic conditions

The first combination is constructed using all the macro factors to reflect relatively complete macroeconomic conditions. This broad category might capture the changes in the business cycle, even though some factors that performed well in the past deteriorate in the present. Previous studies suggest that one group of factors may only mimic changes of macroeconomic conditions in a specific interval (Shiskin and Moore, 1967; Zarnowitz, 1992). These factors are called augmented macroeconomic conditions in Table 2.

(2) Generalized macroeconomic conditions

The second combination is constructed from 106 factors, and selection of these series is motivated by Shiskin and Moore (1967). They propose grouping macro indicators thereby generalizing the strategic processes in the business cycles with macro factors classified into 9 types of economic group: (1) employment and unemployment, (2) production, income, consumption, and trade, (3) fixed capital investment, (4) inventories and inventory investment, (5) prices, costs, and profits, (6) money and credit, (7) foreign trade and payments, (8) federal government activities, and (9) economic activity in other countries. They emphasize that the first two groups are measures of aggregate economic activity and are used to show the broad movements of the business cycle and to determine the dates when there is an economic expansion and contraction starting or ending. He further declares that the next four from fixed capital investment to money and credit with the first two groups can also mirror the business cycle with a causal role in the cyclical process. The last three groups do not contribute to the cyclical fluctuations in U.S.; however, they significantly impact their pattern, amplitude, and duration. Then this study selects majority of the macro factors from the group 1 to group 6. These factors may show generalized macroeconomic conditions, which is shown in Table 2.

(3) Incomplete macroeconomic conditions

This type of macroeconomic condition can be represented by two combinations of macro factors, respectively. One is constructed by using 42 leading factors, shown in Table 2. Leading factors might be expected to move before the business cycle occurs, such as new orders, housing starts, and consumer sentiment. An alternate combination is constructed by pro- and counter- cyclical factors in Table 2. We split these into two parts according to the relation between the business cycle and each economic factor. One part contains 86 procyclical factors and the other covers 28 countercyclical factors. The former group has negative relationship with default risk and the latter positive relationship.

(4) Specific macroeconomic conditions

We introduce a recently advanced variable selection technique, the least absolute shrinkage and selection operator (Lasso) to extract effective variables from historical default patterns. This method is developed has been widely used in economics and finance literature to extract relevant predictive variables and potentially improve prediction accuracy (Tian et al., 2015; Nazemi et al., 2018; Kolari et al., 2020). We use Lasso regression to extract 29 effective factors from 114 macro factors for explaining IDPs.[5] They may move earlier than the other factors in the macroeconomic system which impact on corporate default.

2.2.2 Measure of interactions in the macroeconomic system

The last section shows five types of macroeconomic systems. We measure the interactions among corresponding macroeconomic components to capture the critical transitions in each macroeconomic system, and we finalize macro indicator for each of them. The method of constructing a macro indicator is shown as follows:

$$\text{Indicator} = \frac{\sum_{j>k,|r_{jk}|>0.5}|r_{jk}|}{N} \tag{1}$$

where r_{jk} $(j > k)$ is Pearson's r between macro factors. This method only concerns strong correlation (correlation coefficient larger than 50%) among these macro factors.[6] N represents the number of macro factors in each combination. Equation (1) shows the average of the correlation coefficient in the correlation matrix of each combination of macro factors. We then use rolling-windows, length 9 months, to construct the time series of the indicator. Thus, the value of indicators reveals significant and dynamic relationships between different factors in the macroeconomic system. The presence of a high degree of cross-correlation between the synchronous time evolution of a set of factors shows the risk of unpredictable external shock, and the critical transitions can be captured by examining the fluctuation of the indicators. According to NBER recession date announced by the NBER's Business Cycle Dating Committee, the shortest economic crisis occurred in the beginning of 1980s, which lasted 6 months. The longest occurred in 2007 and lasted 18 months. The length of recession averages about 9 months if extremes are eliminated (Kliesen, 2003).[7] Five indicators based on five combinations of macro factors

---

[5] Prior studies identify that Lasso can select the most significant predictors with strong economic meanings and handle naturally the multi-collinearity problem in bankruptcy/distress prediction, outperforming other selection approaches (Li et al., 2021; Tian et al., 2015). LASSO is able to find a solution for the following model:

$$y = \beta_1 x_1 + \beta_1 x_1 + \cdots + \beta_1 x_1 + \varepsilon$$

by minimizing the prediction error subject to the constraint that the model is not too complex. We apply LASSO to extract key predictors (representing macro factors) for IDPs. Lasso measures complexity by the sum of the absolute values of $\beta_1, \beta_2, \dots, \beta_3$.The solution is obtained by minimizing:

$$\frac{1}{2N}(y - X\beta')'(y - X\beta') + \lambda \sum_{j=1}^{p} |\beta_j|$$

The first term, $(y - X\beta')'(y - X\beta')$, is the in-sample prediction error. It is the same values that least squares minimizes. The second term, $\lambda \sum_{j=1}^{p} |\beta_j|$, is a penalty that increases in value the more complex the model. This term causes lasso to omit variables. they are omitted since the nondifferentiable kinks in the $\sum_{j=1}^{p} |\beta_j|$ absolute terms. The coefficients are shrunk toward 0 correspondingly as λ increases. When λ becomes large sufficiently, some coefficients can be removed insignificant ones from the model with a value of zero, reaching the target of variable selection.

[6] The motivation behind using 50% as a threshold is twofold. First, 50% in the α can keep strong correlation thereby removing some fake correlation among macro factors existing in the weak and moderate apparent correlations (Bretz et al., 2010). In this study, we need to calculate numerous correlation coefficients among different pairs of macro factors. In the statistical tests, there is a problem of Type 1 error, called the multiple comparisons problem (Bretz et al., 2010). In order to avoid this problem, we introduce a method called multiple comparisons. We use two approaches developed by Benjamini and Hochberg (1995) to correct p-values by using adjusted p-values to determine the true correlation coefficient. However, we find that the results are ambiguous, and the indicator is volatile without showing any signs. Since we even adjust p-values, another big issue of whether 5% significance level is good enough has not been handled for statistical scholars. However, we can use 50% to keep strong correlation. Second, even when we use 0% as a threshold to calculate indicators, we find that they have similar results.

[7] NBER's Business Cycle Dating Committee establishes and maintains the chronology of U.S. business cycles and the average recession. The shortest downturn lasted 6 months at the beginning of 1980s, while the longest have lasted 16 months for both 1973-75 and 1981-82. 6-month is quite small for constructing the indicators since it only contains 6 monthly observations for calculating a value of the indicator.

in Table 2 are renamed as indicator_all, indicator_leading, indicator_procyclical, indicator_6_economic, and indicator_effective.

2.3 Summary of statistics

We finalize the sample data for this study including the time series data of IDPs, defaults, and five macro indicators, respectively. Panel A in Table 3 shows summary statistics for these variables. The value of IDP is quite small. The standard deviation of defaults is larger, suggesting that the dispersion is quite high since the largest default month has 37 default events and the smallest default month has no default events. For the macro indicators, each statistic value for both indicator_all and indicator_6_economic is similar, indicating the features of macroeconomic conditions represented by both two indicators are almost similar and they might have similar trend in the past.

3. Statistical Analysis

This study employs several statistical methods to investigate the relationship between IDPs and macro indicators, which is followed by the description of approach used to verify the predictive power of macro indicators on default clustering.

First, we filter IDPs and indicators to remove cyclical components, keeping their trend components. Thus, we smooth the time series as well as the disturbance from zero defaults. The method is based on the Hodrick-Prescott (HP) filter and the penalty parameter we use for smoothing trend is 5.[8] Panel B of Table 3 shows the summary statistics. The statistical properties have parallels in Panel A, which means that the filtering does not change the data structure. More importantly, there is no zero value in IDPs. After filtering variables, we further transforms each variable by building logarithmic ratios so as to transform nonstationary data into stationary ones, following the method of data transformation proposed by Pfaff (2008). Finally, we obtain log-transformed macro-indicators, log (macro indicators), and log-transformed IDPs, log (IDPs).

Second, identifying episodes of structural change that occurred in the periods prior to economic recession helps to explore the sources of recession. We use the breakpoint detection method developed by Bai and Perron (2003) to consider the structural breaks and identify changes in regime in the time series of log (IDPs) before conducting cointegration and causality analysis. Here there are two major concerns. The primary one is to ensure causality analysis is conducted in a stable environment to avoid forecasting errors and unreliability of the model. The second reason is that we can investigate causal relationship between log (IDPs) and log (macro indicators) across different regimes.

The number of structural breaks in each log (macro indicator) is less than those in log(IDPs), and these structural breaks almost form subsets of all structural breaks in the time series of log(IDPs). Thus, we only concern the structural breaks in the time series of log(IDPs) and split the full sample from July 1980 to December 2015. In particular, 4 breakpoints are identified, which occurred in October 1985, June 1993, November 1998 and December 2003. The best combination of breakpoints is confirmed by BIC criterion. Finally, we have five subsamples, namely, sample 1980.09 – 1985.09, sample 1985.10 – 1993.06, sample 1993.07 – 1998.11, sample 1998.12 – 2003.12 and sample 2004.01 – 2014.12. As shown

As a robustness test, we also used 12 months and 18 months to construct indicators. We find that the indicators constructed by using from 12 to 18 months tend to be Granger-caused by corporate defaults and these indicators from 6 to 18 months have a decreased tendency of Granger-causing corporate defaults. That is, the predictive effects of macroeconomic conditions on corporate defaults gradually reduces, and corporate default gradually has more predictive information for macroeconomic conditions. These results correspond to intuition. From a statistical perspective, when the window size is expanded, more information is covered in the calculation of an indicator. This may bring more opaque information to the calculator and the indicator is not able effectively to reflect changes in macroeconomic conditions. The results can be provided upon the request.

[8] We also used penalty parameter from 1 to 4. The results are similar.

in Panel F of Figure 2, fundamental structural changes in corporate default are associated with recessions. While the third sample (1993.07–1998.11) does not cover a period of recession, all other samples include the occurrence of economic recessions.

Third, we use a cointegration test proposed by Johansen (1988) to investigate whether there is a long-run relationship between each log(macro indicators) and log(IDPs) in these abovementioned samples. This method can also provide a cross-check on the validity of the results at the end of the causality analysis. The results are shown in Table 5. We further employ two types of causality methods to investigate the causal relationship between each log (macro indicator) and log (IDPs) in different samples. Based on the results from unit root tests in Table A1 in the appendix, we identify whether these series are stationary or not. If both variables are stable, we use Granger causality to examine a bidirectional relationship between each log(macro indicators) and log(IDPs); otherwise, we use Toda and Yamamoto causality test proposed by Toda and Yamamoto (1995). In addition, the diagnostics tests for these models between each log (macro indicator) and log(IDPs) in Table A2 shows whether causality models pass the diagnostics tests or not.

Fourth, we use a linear regression model to statistically verify whether five log transformed macro indicators can predict default clustering. Here, default clustering can be measured as follows. Basically, we use the monthly default events equal to 5 as a filtering standard to remove small default events and obtain clustered defaults. In order to make a distinction between original IDPs and filtered IDPs removing small default events, we use FIDPs instead as default clustering. The primary reason is that the mean of the number of defaulters in Table 3 is 5.445 and the value of 25% of total default sample is 2. Removing the months with less than 5 defaults can statistically keep small default events removed. Interestingly, we find there is a rough linear relationship between log(FIDPs) and each log(macro indicators) via scatter plot analysis if the months having less than and equal to 5 default events are deleted.[9] Additionally, using 5 can lead to best performance in the regression model; that is, predicting power of each indicator reaches the highest level compared with using the other values for FIDPs.[10] Also, the fitted model using 5 can satisfy model assumptions. We use 5 to filter small defaults to identify default clustering. After removing small defaults, the number of observations is reduced from 411 to 169.

Figure 3 shows the time series of FIDPs and macro indicators after removing corporate default smaller than and equal to 5. For model construction, we also consider the impact of economic recession on default clustering. A dummy variable representing NBER economic recession indicator as well as the cross term between recession indicator and log (macro indicator) are added into the model, which can give a full picture of how our log (macro indicator) representing macroeconomic conditions impact on default clustering. Here, the regression analysis is based on two approaches. One is to use each indicator and NBER economic recessions without any lags; the other is to further verify whether both macro indicators and NBER economic recessions have predictive ability for FIDPs 3 months ahead.

4. Empirical Results

The main empirical results are presented in this section, including changes of macro indicators and IDPs over time, the relationship between macro indicators and IDPs and the prediction of default clustering using macro indicators.

4.1 Dynamics of macro indicators and IDPs

Figure 2 shows monthly fluctuations in each macro indicator and IDPs. From Panel A to Panel E, the indicators are indicator_all, indicator_leading, indicator_procyclical, indicator_6_economic, and

[9] The results can be provided upon the request.

[10] We use R-squared and adjusted R-squared to measure the explanatory power for each macro indicator on FIDPs.

indicator_effective, representing different types of macroeconomic conditions. All the indicators are blue line in each Panel and IDPs are the red line. The green bars represent four historical NBER economic recessions.

In each panel, the blue line representing macro indicator have a similar trend in IDPs except two main intervals including the beginning of 1980s and 2000s, suggesting that correlation among macro factors signals the future trend of corporate default or measure corporate default. For the two subsamples 1980.09 – 1985.09 and 1998.12 – 2003.12, this is not the case. The US economy experienced economic recessions. The first occurred in January 1980 and lasted 6 months before the economy returned to growth. A year later, the economy fell again. This recession lasted 1 year and 4 months. Both recessions were triggered by higher inflation and an oil crisis. In particular, monetary policy implemented by Federal Reserve contributed to double-digit inflation in 1980. After that, the Federal Reserve undertook to tighten monetary policy in order to curb inflation, which led the economy into a deep recession (Mishkin, 1995). During this period, many economic sectors, such as the manufacturing and construction sectors, suffered from higher defaults among firms. In the beginning of 1980s few defaults are recorded by Moody's; this is also identified by previous studies (Lando and Nielsen, 2010).

All the indicators are quite volatile in the beginning of the 1980s. After the 1980s recession, majority indicators experience a higher increase, suggesting that many factors have higher correlation among each other, and macroeconomic conditions are still poor. This phenomenon is noted in the economic literature. For example, Rattner (1981) reports that "top officials at the Federal Reserve Board, including its chairman, Paul A. Volcker, say that their policy of reducing the expansion of money and credit will mean little or no economic growth in 1981 and continuing high interest rates". After failing to gain traction during the weak and brief recovery from the 1980 downturn, weakness in manufacturing and housing caused by rising interest rates began to have an expanded effect on related sectors beginning in mid-1981 (Bednarzik et al., 1982). Higher interest rates lead to the other macro factors moving then. Therefore, the correlation among these macro factors is high. These arguments provide the reason why there is a high co-movement among macro factors or why all the macro indicators are quite high after July 1981.

In the 2000s recession, the peak points in each indicator are lag of the peak point in the series of IDPs. It is suggested that higher correlation among macro factors reach the peak level later compared with peak of corporate default. Interestingly, IDPs reach a peak in the middle of the 2000s recession and then decrease rapidly. In contrast, IDPs experience rapid increase in the 1990s and 2007-2009 financial crisis until the official ending point of recession and then they decrease fast in the non-recession period.

As mentioned above, although five macro indicators have similar trend with IDPs except during two special periods including early of 1980 and 2001, they show different trends in certain intervals. The peaks of some indicators are ahead of peak IDPs in certain periods. For example, in Panel A of Figure 2, the blue line representing indicator_all peak in the official NBER cyclical trough in March 1991, which is the same as for the IDPs. While the peak points in the indicators constructed by the other types of factors in Panel B, C, D, and E follow that in the IDPs. Interestingly, both all the indicators and IDPs peak in the formal end to the recession in June 2009. In contrast, IDPs have a tendency of being ahead of macro indicators in the certain periods except the recession of 2001.

The trend of indicator_all is similar to the trend of the indicator_6_economic, indicating that the co-movement among these factors from 6 economic groups account for the large weight in the co-movement among all the factors. More interestingly, in Panel E, the indicator_effective seemingly displays a similar trend of IDPs, suggesting that historical default crisis can be reflected by dynamic of correlation among certain factors. This result corresponds to the intuition since Lasso is used to extract effective predictors for IDPs and then these predictors should have the capability to predict IDPs.

In this context, we further introduce causality analysis to explore whether there is a causal relationship between IDPs and each macro indicator thereby identifying causal relationship between changes in macroeconomic conditions and corporate default.

4.2 Long-run relationship between macro indicators and IDPs

Based on the results of the unit root test in Table A1, we confirm which sample IDPs and indicator may be cointegrated in Table 4. It shows 10 samples that may have cointegration between IDPs and macro indicators. Table 5 describes the cointegration tests. The test statistics and asymptotic 5% critical values are shown in Panel A and Panel B. There are just 3 out of 30 subsamples that have cointegration between IDPs and macro indicators constructed by factors from 6 economic groups and effective factors.

Both tests reject the hypothesis of no cointegration ($r = 0$) at the 5% level, whereas they do not reject the hypothesis that $r <= 1$. Therefore, the conclusion is that $r = 1$, that is, there is one stationary relationship between the level of the variables. Since the indicator_6_economic is able to reflect cyclical cycle as using all the factors, the dynamic of the entire macroeconomic conditions tends to be cointegrated with corporate default risk in the long run. Another indicator cointegrated with IDPs is the indicator_effective. This result corresponds to intuition since the effective factors can explain well historical default from January 1980 to December 2014 and this indicator should relate with IDPs.

The cointegrated relationship between IDPs and two macro indicators constructed by factors in 6 economic groups and effective factors implies causality in at least one direction. Then causality tests should be used to further determine the causal relationship between them.

4.3 Causal relationship between macro indicators and IDPs

Table 6 presents the results of Granger-causality and TY causality tests. Panel A shows causality results between indicator_all and IDPs. In the full sample, we find bidirectional causality between indicator_all and IDPs since this indicator (IDPs) is significant at 1% (10%) significance level. There is no causal relationship in the subsample from December 1998 to December 2003 which are not significant at 5% or even 10% significance level. In contrast, in two subsamples including September 1980 – September 1985 and October 1985 – June 1993, we reject that IDPs does not Granger-cause indicator based on a significance level of 1% and a significance level of 5%. It is indicated that IDPs can provide predictive content of indicator_all. However, indicator_all can Granger-cause IDPs in the two samples including July 1993 – November 1998 and January 2004 – December 2014 since the p-values are highly significant at 5% and 1% significance level.

Panel B shows the results of using indicator_leading. In the full sample, in terms of Granger causality, causality flows from the indicator to IDPs, but not from IDPs to the indicator. There are only three subsamples that have directional causal relationship between indicator and IDPs. For example, in the first interval from September 1980 to September 1985, indicator_leading can Granger-cause IDPs; in the next interval, the causal relationship reverses and IDPs can provide predicting information for this indicator. The last interval shows indicator_leading Granger-causes IDPs again with statistical significance. This panel thus indicates that there is a tendency of directional causal relationship running from the indicator to IDPs.

Panel C shows the causality results between IDPs and indicator_procyclical. Likewise, this indicator can provide predictive information for IDPs in the full sample and the last interval of the full sample, which is same as two indicators mentioned above. In the subsample from September 1980 to September 1985, there is a bidirectional relationship between indicator and IDP since both two p-values are significant, suggesting that the indicator and IDPs Granger-cause each other. Interestingly, in the period from July 1993 to November 1998, IDPs are significant at the 5% level in terms of Granger-causing this indicator.

Panel D provides causality results after using indicator_6_economic for generalized macroeconomic conditions. The causal relationship results are the same as from using indicator_all except two points.

One is that the indicator can Granger-cause IDPs if the significance level of 1% is used for rejecting the hypothesis of no causal relationship in the full sample; another is that there is no directional relationship between IDPs and indicator in the period from July 1993 to November 1998. It can be concluded that the results of using indicator_all are similar to that of using indicator_6_economic. This suggests that generalized macroeconomic conditions are able to reflect augmented macroeconomic conditions based on all the factors. This might have been expected, since the number of factors from the 6 economic groups (104) is only 10 less than the total number of factors (114).

Panel E is a special case compared with the others. First, the indicator is constructed by using only 29 effective factors, which can explain historical default. Second, this indicator has the most predictive information on explaining IDPs since it can Granger-cause IDPs in the four out of 6 subsamples with highly significant p-values. Third, IDPs have weak ability of Granger-causing this indicator.

This result corresponds to our intuition. These 29 factors are the key elements in terms of explaining historical defaults. In addition, based on the theory of anticipating critical transition, these components in the macroeconomic system initially move before the other components. Therefore, following the suggestions from Helbing (2013), if we can capture the interaction among these factors and excluding the opaque interaction information from the other factors, we can get more efficient indicator to explain corporate defaults happening in the past 35 years. Clearly this indicator should have most predictive power on IDPs.

There are several notable outcomes:

1) Panels A and D show similar findings. First, IDPs Grange-cause indicator_all and indicator_6_economic in the two subsamples including 1980.09-1985.09 and 1985.10-1993.06. These results provide empirical evidence support of how deteriorated credit market conditions have impact on the economy. Interestingly, Bernanke (1993) argues that the savings and loans crisis in the 1980s and 1990s is a good example of corporate default making an economy worse. Likewise, Jermann and Quadrini (2012) argue that the changes in credit conditions resulted in the economic downturns in 1990s.

2) Both two indicators mentioned above Granger-cause IDPs in the subsample from January 2004 to December 2014, implying that macroeconomic conditions contain predictive information on corporate default. This finding corresponds to the theoretical studies of Hackbarth et al. (2006) who argue that a default threshold is associated with tax benefit and bankruptcy costs. The former depends on the level of cash flow, which in turn depends on macroeconomic conditions since in boom times cash flows tend to be higher and in busts lower. The latter is dependent on the probability of default and the loss given default, which all depend on the current macroeconomic conditions. Then macroeconomic conditions can influence corporate defaults. This is consistent with Bhamra et al. (2011) who declare that the financial crisis of 2007-2008 is an example of these changes having a severe impact on both default rates and credit spreads of firms. In contrast, there is no causal relationship between IDPs and two indicators in the subsample from December 1998 to December 2003.

3) In panels B and E, IDPs have less predictive information for the indictors constructed by using leading factors and effective factors. In contrast, two indicators have more predictive information for IDPs. These results can be explained below. Leading factors tend to change the direction earlier than the business cycle, and effective factors can explain what happened in corporate default.

4) Although each indicator can all Granger-cause IDPs from September 1980 to December 2014, this finding is not reliable since these models do not pass the model diagnostic tests in Table

A2. In contrast, all the models in the subsamples for each indicator pass the model diagnostic tests.

4.3 Investigation of whether and how macro indicators explain FIDPs

This section shows the results of investigating whether and how macro indicators explain FIDPs.

4.3.1 First approach: no lags in macro indicator and economic recession indicator

Results are in Table 7. Five macro indicators and the recession indicator are all significant in explaining FIDPs, suggesting that both changes of macroeconomic conditions and economic recessions can explain higher corporate default. The interaction items defined as indicator times recession are significant in using 5 macro indicators and are all negative. This means that, during recession, changes of a unit of indicator should have less effect on FIDPs compared with non-recession. That is, changes of defined macroeconomic conditions have different effects on default clustering in recession and non-recession.

Next, we investigate the unique effect of each indicator on default clustering. we find the effective indicator outperform others in both economic and statistical sense. Specifically, besides higher value of R^2, we find that FIDPs' sensitivity to changes in the effective indicator is higher than its sensitivity to other indicators. As shown in Panel B of Table 7, if there is no recession, the indicator_effective has the largest unique effect on FIDPs with coefficient 0.62, followed by indicator_all (0.616), indicator_6_economic (0.597), indicator_procyclical (0.558), and indicator_leading (0.537).

In recession, the unique effect of each indicator on FIDPs is the sum of $\beta_1 + \beta_3$ shown in Panel C. Then for the indicator constructed by all the factors, leading factors, procyclical factors, factors in 6 economic groups and effective factors, the unique effects are 0.053, 0.174, 0.036, 0.056 and 0.05 respectively. Clearly, the unique effect of each indicator on FIDPs in the recession are notably smaller than their effects on FIDPs in non-recession. That is, a rapid change of macro indicators and recession indicator has larger effect in explaining clustered default in the non-recession period but a smaller effect on predicting clustered default, which is consistent with findings in Kim et al. (2017).

This finding can be interpreted as below. The NBER's Business Cycle Dating Committee defines NBER recession. The committee examines and compares the behavior of various measures of economic activity: real GDP measured on the product and income sides, economy-wide employment, and real income. They also may consider indicators that do not cover the entire economy, such as real sales and the Federal Reserve's index of industrial production. If these indexes decrease consecutively, NBER recession is defined. Then economic recession is a special scenario in the evolution of economic environment. Figure 3 shows that macro indicators have experienced rapid increase in each recession except the recession occurring in the early 1980s. While the peaks of FIDPs exist during the recession period, it can be expected that there is an interaction between NBER recessions and macro indicators.

In this context, changes in macroeconomic conditions, as reflected by changes in interest rates, the stock market indexes, exchange rates, unemployment rates, etc. may impact the overall profitability of firms. The exposures increase of the probabilities of default and of migration from one credit rating to another (Crouhy et al., 2000) and corporate default increases during the period. The indicator_effective, recession and their interaction item have the largest predicting power on FIDP with adjusted R-squared being equal to 38%, which is followed by using indicators constructed by indicator_all (36%), indicator_6_economic (35%), indicator_leading (32%), and indicator_procyclical (30%).

Figure 4 lists the results of model diagnostics. Panel A shows two sets of information; one for the residual and fitted plot, the other as a QQ plot for the normality test in the residual. For the residual vs fits plot, this is a scatter plot of residuals on the y axis and fitted values (estimated responses) on the x axis. Majority plots illustrate that the scatter points seemingly cluster in the right side of each graph. However, it seems that the residuals randomly distribute around the 0 line especially for Panel A1, Panel A4, and Panel A5. The assumption of a linear relationship seems to be reasonable. In addition, the residuals do not form a horizontal band around the 0 line. We use a heteroscedasticity test to decide whether the variances of the error terms are constant or not. Since all the p-values are larger than 10%, the null hypothesis of constant in the error term is not rejected. Thus, the variances of the error terms in each model are equal. A QQ plot is used to test whether the residuals follow a normal distribution. The points on the QQ plot fall approximately on a straight line in Panel A1, Panel A3, and Panel A4. Two cases are not so good: in Panel A2 and Panel A5, since there are several points beyond the 5% area. That is, we may reject the null hypothesis of normality at the significance level 5%. From these results, we can conclude that majority models are fitted well. Particularly, both indicator_all and indicator_6_economic are well fitted compared with the others.

Thus, both changes of macro indicators and recession can provide predictive content on default clustering. There is an additive effect between them. During recession, the effect of different types of macro indicators on default clustering is smaller than that in non-recession. The indicator_effective, indicator_all and indicator_6_economic have the largest predicting power with recession and their interaction item. From results of causality tests, these three indicators provide more predictive content than the other types of indicators. That is, these results correspond to the findings from causality analysis since they have more causal relationships in different intervals. Specific macroeconomic conditions outperform non-specified one, even the augmented and generalized macroeconomic environment.

4.3.2 Second approach: lagged macro indicators and economic recession indicators by 3 months

Table 8 shows the regression results after lagging indicator and NBER recession by 3 months. Compared with the results in Panel A in Table 7, the findings are almost same. Specifically, five macro indicators and recession are still significant in explaining FIDPs, suggesting that both changes of various macroeconomic conditions and recessions can predict default clustering 3 months ahead.

In Panel A, although both NBER recessions and indicators can still predict FIDPs, it is not the same case for the interaction items between them. None is statistically significant, which means that there is no interaction effect between lagged indicators and lagged NBER recession.

For the results of adjusted R-squared, the largest is still from using the indicator_effective. Interestingly, the second-best indicator is indicator_leading, and its adjusted R-squared increase from 0.319 in Table 7 to 0.331 in Table 8. In contrast, the other corresponding values of adjusted R-squared decrease, for example, adjusted R-squared for the indicator_effective reduce from 0.383 to 0.332 by 0.051. The value of adjusted R-squared constructed by all the factors goes down to 0.321 by 0.033 from 0.354. Likewise, for the indicator_6_economic, adjusted R-squared value decreases from 0.346 to 0.317 by 0.029.

Figure 5 provides the diagnostics results of models in Table 8. For the residual vs fits plot, most plots show the scatter points somewhat clustered in the right side of each graph. However, it seems that the residuals randomly distribute around the 0 line. Thus, it may be concluded that the assumption of a linear relationship seems to be reasonable. In addition, the residuals do not form a horizontal band around the 0 line since the blue line in each graph are not purely along the zero line in the left side of each graph. Heteroscedasticity is tested to see whether the residual is constant. Since all the p-values are larger than 5%, the null hypothesis of constant in the error term is not rejected and the variances of the error terms in each model are equal. A QQ plot is used to test whether the residuals follow normal distribution. It can be seen that all the points on the QQ plot fall approximately on a straight line. We can conclude that majority models are fitted well.

Both changes of each indicator and recession can provide predictive content on default clustering by 3 months ahead. The additive effect disappears after lagging indicator and economic recession by 3 months. That is, the unique effect of changes of each indicator on default clustering is constant in both recession and non-recession. The indicator_effective have strongest predicting power, which is followed by indicator_leading, indicator_all, indicator_6_economic, and indicator_procyclical. Incomplete macroeconomic conditions still have explanatory for default clustering compared with specific and augmented and generalized macroeconomic conditions. All in all, these empirical results provide supporting evidence for theoretical studies including Chen (2010) and Hackbarth et al. (2006). That is, macroeconomic conditions can make corporate default worse and even lead to default clustering. Even they have 3-month's predictive effect on default clustering.

5. Discussion

We split our discussion into two parts: consideration of the interaction of macroeconomic conditions with corporate default and the possibility of building an early-earning system for capturing default clustering.

5.1 How do macroeconomic conditions interact with corporate default?

Macroeconomic conditions interact with corporate default in different intervals, reflecting the existence of feedback causality.

On one hand, corporate default can impact macroeconomic conditions. Credit market frictions result in the changes of credit market conditions, measured by firm defaults, and deteriorated credit market conditions can depress economic activity. For example, from September 1980 to June 1993 corporate default exerted a powerful effect on the economy, which led to the economy slowing down and macroeconomic conditions becoming worse. This period covers two economic recessions, which are the savings and loans crisis in the 1980s and 1990s. As Bernanke (1993) argues, federal deposit insurance provided extensive credit to the savings-and-loan institutions without enforcing sufficient limits on the riskiness of savings-and-loan investments. Saving-and-loan owners were motivated to engage in highly levered and risky portfolios of long-term loans, mortgage-backed securities, and other risky assets.

On the other hand, macroeconomic conditions can impact corporate default. Corporate defaults are treated as endogenous events in default structural models (Black and Cox, 1976; Leland and Toft, 1996; Acharya and Carpenter, 2002), in which corporate defaults depend on shareholders' default policy represented by various default thresholds for each economic state. Once debt has been issued, shareholders have the option to default on their obligations. Hackbarth et al. (2006) explain that a default threshold is strongly associated with two main elements including tax benefit and bankruptcy costs. The former depends on the level of cash flow, which in turn depends on macroeconomic conditions. The latter is dependent on the probability of default and the loss given default, which all depend on the current macroeconomic conditions. Several studies find that changes of monetary policy representing macroeconomic conditions have strong impact on corporate leverage decisions and defaults via its effect on inflation or deflation (Wadhwani, 1986; Bhamra et al., 2011; Narayan et al., 2021). Then macroeconomic conditions can have a great influence on corporate defaults. We find in the two intervals including July 1993 – November 1998 and January 2004 – December 2014, macroeconomic conditions can affect corporate default. Particularly, the latter interval covers the financial crisis of 2007-2008. This is consistent with Xiao et al. (2009) and Bhamra et al. (2011).

In addition, the causal relationship between macroeconomic conditions and corporate default is insignificant in certain phases of business cycles. In this study, we find that this is the case in the interval covering the recession occurring in March 2001. In fact, the economic literature has already explained this finding (Kliesen, 2003), where the argument is that this recession had a unique feature whereby the economy experienced substantial growth - four monthly coincident indicators used by the NBER Business Cycle Dating Committee to establish its business cycle chronology were growing by December

2001. [11] We see that there is a conflict between good economic environment and default clustering in that period. Then the causal direction from macroeconomic conditions to corporate default is not logically reasonable. In addition, this recession is caused by the firms that cut back on expenditures in the information technology (IT bubble) rather than shopping mall and corridors of the Federal Reserve Bank (Stock and Watson, 2003). That is, the problem is not due to economic structure but a specific industry, which has minor influence on whole economy (Kliesen, 2003). Intuitively, capital intensive and consumer industries are highly related to aggregate demand, output, and employment, since they mainly contribute to the fluctuation on demand, output and employment. However, technology firms are not the same case; for example, they do not need to employ large numbers of staff and establish numerous factories with much equipment and facilities.

5.2 Can we construct an early-earning system for capturing default clustering?

According to the results shown in Table 7 and Table 8, we can conclude that macro indicators representing macroeconomic conditions have exploratory ability on FIDPs representing default clustering. That is, macroeconomic conditions can make corporate default worse thereby leading to default clustering, which is consistent with the theoretical studies of Chen (2010) and Hackbarth et al. (2006). Interestingly, they also have predictive power on default clustering 3-month ahead.

The best indicator for predicting default clustering is constructed with a set of 29 effective factors. Additionally, the leading indicator has the best predictive power during recession. This suggest that the leading macroeconomic factors move at an earlier stage than the other factors such as lagging and coincident factors and they tend to change direction before the business cycle occurs. This verifies the evidence that factors of recession prediction can vary over business cycle (Hwang, 2019). The timing difference between the cyclical turn in the indicators and business cycle turn is one month or more, such as, new orders, housing starts, and consumer sentiment. Then the interaction among them can give early sign for the macroeconomic system. This indicator performs well after lagging by 3 months.

We also find that the indicators constructed by using more macroeconomic information including all the factors, factors from 6 economic groups, and procyclical factors respectively, have lower predictive power in comparison with two indicators constructed by using fewer factors. This finding gives empirical support for Helbing (2013), who claims "too much data can make it difficult to separate reliable from ambiguous or incorrect information, leading to misinformed decision-making. Hence too much information may create an opaquer rather than a more transparent picture."

To sum up, these empirical findings suggest that we could build an early-warning system for predicting default clustering based on these macro indicators especially for the indicators constructed by effective factors as well as leading factors.[12]

[11] The monthly indicators are employment, industrial production, personal income and manufacturing and trade sales, which are used by NBER Business Cycle Dating Committee to establish its business cycle chronology. Specifically, real GDP even increased 0.2 percent from the first quarter of 2001 to the fourth quarter of 2001. Nonfarm labor productivity increased 2.2 %, which is more than a percentage point faster than during the average postwar recession. Strong labor productivity growth also helped to keep real disposable personal income growth positive during the recession (0.37 %), rather than declining slightly as is typically the case. Hence, helping to underpin the strength of real consumer spending during the recession was relatively strong growth of nonfarm labor productivity.

[12] In fact, these findings in this study also corresponds to the study done by Xing and Yang (2020) who construct the indicators by using 6-month lagged data with a similar equation as this study's approach to predict corporate default and confirm that an indicator constructed by leading factors does have predictive information for corporate default.

6. Conclusion

This paper constructs macro indicators to measure macroeconomic conditions in US based on the theory of capturing critical transitions in the complex system. This paper uses these indicators to investigate the causal relationship between macroeconomic conditions and corporate default and examines the predictive power of macroeconomic conditions on default clustering.

In particular, we empirically verify interactions between macroeconomic conditions and corporate default clustering across different economic regimes. Clearly caution is required before macroeconomic factors to predict corporate default since these do not always contain predictive information on corporate default. Additionally, we identify macroeconomic conditions can predict default clustering. Our results are consistent with recent theoretical findings of uncertain macroeconomic environment can lead to default clustering. Meanwhile, we find that the best three indicators are constructed by using effective factors, leading factors, and all the factors respectively in terms of prediction of default clustering. This implies the possibility of building an early-warning system capturing clustered defaults.

For future study, we can build an early-warning system for clustered default events in practice. Based on our study's findings, we may use three types of macro indicators. One is the macro indicator constructed by effective indicator extracted by Lasso; another is the macro indictor constructed by leading factors and all the factors used in this study. For the former one, we may extract effective factors based on the changes of the time-length for prediction. This can better reflect the major components in the real-time macroeconomic system, and our approach can punctually capture the sudden critical transitions in the macroeconomic system. The three approaches are all used for forecasting clustered defaulters.

Moveover, it is worth further exploration how to build more effective early-warning system for predicting corporate default by constructing new indicators. For example, Figure in the Appendix shows the risks interconnection map 2011 illustrating systemic interdependence in the hyper-connected world, taken from Helbing (2013). The credit crunch/liquidity that is used to measure credit market conditions has five main direct connections with asset price collapse, fiscal crisis, global imbalances and currency volatility, extreme consumer price volatility, and regulatory failures. However, these five terms are highly correlated with the other types of risks. Therefore, how to construct default risk indicator by effectively and efficiently using various data, such as macro factors, micro factors, political factors, and environmental factors is the next challenge in term of predicting future collapse in the credit market.

References

ACHARYA, V. V. & CARPENTER, J. N. 2002. Corporate bond valuation and hedging with stochastic interest rates and endogenous bankruptcy. *Review of Financial Studies,* 15**,** 1355-1383.

ALTMAN, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance,* 23**,** 589-609.

AZIZPOUR, S., GIESECKE, K. & SCHWENKLER, G. 2018. Exploring the sources of default clustering. *Journal of Financial Economics*, *129*(1), 154-183.

BAI, J. & PERRON, P. 2003. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics,* 18**,** 1-22.

BATTISTON, S., FARMER, J. D., FLACHE, A., GARLASCHELLI, D., HALDANE, A. G., HEESTERBEEK, H., HOMMES, C., JAEGER, C., MAY, R. & SCHEFFER, M. 2016. Complexity theory and financial regulation. *Science,* 351**,** 818-819.

BEDNARZIK, R. W., HEWSON, M. A. & URQUHART, M. A. 1982. The employment situation in 1981: new recession takes its toll. *Monthly Labor Review***,** 3-14.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)***,** 289-300.

BERNANKE, B. S. 1993. Credit in the macroeconomy. *Quarterly Review-Federal Reserve Bank of New York,* 18**,** 50-50.

BERNANKE, B. S., GERTLER, M. & GILCHRIST, S. 1999. The financial accelerator in a quantitative business cycle framework. *Handbook of Macroeconomics,* 1**,** 1341-1393.

BHAMRA, H. S., FISHER, A. J. & KUEHN, L.-A. 2011. Monetary policy and corporate default. *Journal of Monetary Economics,* 58**,** 480-494.

BHARATH, S. T. & SHUMWAY, T. 2008. Forecasting default with the merton distance to default model. *Review of Financial Studies,* 21**,** 1339-1369.

BLACK, F. & COX, J. C. 1976. Valuing corporate securities: some effects of bond indenture provisions. *Journal of Finance,* 31**,** 351-367.

BLACK, F. & SCHOLES, M. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy,* 81**,** 637-654.

BOSE, U., FILOMENI, S., & MALLICJ, S. 2021. Does bankruptcy law improve the fate of distressed firms? The role of credit channels. *Journal of Corporate Finance, 68*, 101836.

BRETZ, F., HOTHORN, T. & WESTFALL, P. 2010. *Multiple comparisons using R*, CRC Press.

CAMPBELL, J. Y., HILSCHER, J. & SZILAGYI, J. 2008. In search of distress risk. *Journal of Finance,* 63**,** 2899-2939.

CHAVA, S. & JARROW, R. A. 2004. Bankruptcy prediction with industry effects. *Review of Finance,* 8**,** 537-569.CHEN, H. 2010. Macroeconomic conditions and the puzzles of credit spreads and capital structure. *Journal of Finance,* 65**,** 2171-2212.

COLLIN ‐ DUFRESNE, P., GOLDSTEIN, R. S. & MARTIN, J. S. 2001. The determinants of credit spread changes. *Journal of Finance,* 56**,** 2177-2207.

CROUHY, M., GALAI, D. & MARK, R. 2000. A comparative analysis of current credit risk models. *Journal of Banking and Finance,* 24**,** 59-117.

DAS, S. R., DUFFIE, D., KAPADIA, N. & SAITA, L. 2007. Common failings: how corporate defaults are correlated. *Journal of Finance,* 62**,** 93-117.

DIKS, C., HOMMES, C., & WANG, J. 2019. Critical slowing down as an early warning signal for financial crises?. *Empirical Economics*, *57*(4), 1201-1228.

DUFFIE, D., ECKNER, A., HOREL, G. & SAITA, L. 2009. Frailty correlated default. *Journal of Finance,* 64**,** 2089-2123.

DUFFIE, D., SAITA, L. & WANG, K. 2007. Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics,* 83**,** 635-665.

EOM, Y. H., HELWEGE, J. & HUANG, J.-Z. 2004. Structural Models of Corporate Bond Pricing: an Empirical Analysis. *Review of Financial Studies,* 17**,** 499-544.

FISCHER, E. O., HEINKEL, R. & ZECHNER, J. 1989. Dynamic capital structure choice: theory and tests. *Journal of Finance,* 44**,** 19-40.

GIESECKE, K., LONGSTAFF, F. A., SCHAEFER, S. & STREBULAEV, I. 2011. Corporate bond default risk: a 150-year perspective. *Journal of Financial Economics*, 102**,** 233-250.

GOLDSTEIN, R., JU, N. & LELAND, H. 2001. An EBIT‐Based model of dynamic capital structure. *Journal of Business*, 74**,** 483-512.

GOURIEROUX, C., MONFORT, A., MOUBABBI, S., & RENNE, J. P. 2021. Disastrous defaults. *Review of Finance*, *25*(6), 1727-1772.

HACKBARTH, D., MIAO, J. & MORELLEC, E. 2006. Capital structure, credit risk, and macroeconomic conditions. *Journal of Financial Economics,* 82**,** 519-550.

HELBING, D. 2013. Globally networked risks and how to respond. *Nature*, 497**,** 51-59.

HOMMES, C. 2021. Behavioral and experimental macroeconomics and policy analysis: A complex systems approach. *Journal of Economic Literature*, *59*(1), 149-219.

HWANG, Y. 2019. Forecasting recessions with time-varying models. *Journal of Macroeconomics*, *62*, 103153.

JACOBSON, T., LINDÉ, J. & ROSZBACH, K. 2013. Firm default and aggregate fluctuations. *Journal of the European Economic Association,* 11**,** 945-972.

JARROW, R. & TURNBULL, S. 1992. Credit risk: drawing the analogy. *Risk Magazine,* 5**,** 63-70.

JERMANN, U. & QUADRINI, V. 2012. Macroeconomic effects of financial shocks. *American Economic Review,* 102**,** 238-271.

JOHANSEN, S. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control,* 12**,** 231-254.

JORDAN, B. D. & JORDAN, S. D. 1997. Special repo rates: An empirical analysis. *Journal of Finance,* 52**,** 2051-2072.

KIM, T. S., PARK, J. W., & PARK, Y. J. 2017. Macroeconomic conditions and credit default swap spread changes. *Journal of Futures Markets*, *37*(8), 766-802.

KLIESEN, K. L. 2003. The 2001 recession: how was it different and what developments may have caused it? *Review Federal Reserve Bank of Saint Louis,* 85**,** 23-38.

KOLARI, J. W., LÓPEZ-ITURRIAGA, F., & SANZ, I. P. 2020. Measuring systemic risk in the US Banking system. *Economic Modelling,* 91, 646-658.

KOOPMAN, S. J., KRÄUSSL, R., LUCAS, A. & MONTEIRO, A. B. 2009. Credit cycles and macro fundamentals. *Journal of Empirical Finance,* 16**,** 42-54.

KOOPMAN, S. J. & LUCAS, A. 2005. Business and default cycles for credit risk. *Journal of Applied Econometrics,* 20**,** 311-323.

KOOPMAN, S. J., LUCAS, A. & SCHWAAB, B. 2011. Modeling frailty-correlated defaults using many macroeconomic covariates. *Journal of Econometrics,* 162**,** 312-325.

LANDO, D. & NIELSEN, M. S. 2010. Correlation in corporate defaults: contagion or conditional independence? *Journal of Financial Intermediation,* 19**,** 355-372.

LELAND, H. E. 2004. Predictions of default probabilities in structural models of debt. *Journal of Investment Management,* 2**,** 5-20.

LELAND, H. E. & TOFT, K. B. 1996. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance,* 51**,** 987-1019.

LONGSTAFF, F. A. & SCHWARTZ, E. S. 1995. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance,* 50**,** 789-819.

LÜTKEPOHL, H. 2005. *New introduction to multiple time series analysis*, Springer Science & Business Media.

MCNEIL, A. J., FREY, R. & EMBRECHTS, P. 2010. *Quantitative risk management: concepts, techniques, and tools*, Princeton university press.

MERTON, R. C. 1974. On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance,* 29**,** 449-470.

MISHKIN, F. S. 1995. Symposium on the monetary transmission mechanism. *Journal of Economic Perspectives,* 9**,** 3-10.

MOODY, S. 2015. Moody's rating symbols and definitions. *Moody's Investors Service, Report,* 79004**,** 46.

NARAYAN, S., BUI, M. N. T., REN, Y., & MA, C. 2021. Macroeconomic determinants of US corporate leverage. *Economic Modelling,* 104, 105646.

NAZEMI, A., & FABOZZI, F. J. 2018. Macroeconomic variable selection for creditor recovery rates. *Journal of Banking & Finance*, 89, 14-25.

OHLSON, J. A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research,* 18**,** 109-131.

PFAFF, B. 2008. *Analysis of integrated and cointegrated time series with R,* Springer New York.

SCHEFFER, M., CARPENTER, S. R., LENTON, T. M., BASCOMPTE, J., BROCK, W., DAKOS, V., VAN DE KOPPEL, J., VAN DE LEEMPUT, I. A., LEVIN, S. A. & VAN NES, E. H. 2012. Anticipating critical transitions. *Science,* 338**,** 344-348.

SHISKIN, J. & MOORE, G. 1967. Indicators of business expansions and contractions. *Occasional Paper*, 103.

SHUMWAY, T. 2001. Forecasting bankruptcy more accurately: a simple hazard model. *Journal of Business,* 74**,** 101-124.

SORNETTE, D. & WOODARD, R. 2010. *Financial bubbles, real estate bubbles, derivative bubbles, and the financial and economic crisis*, Springer.

STEVEN RATTNER. 1981. *Fedral reserve sees little growth in '81 with contitinued high rates* [Online]. New York: New York Times. Available: http://www.nytimes.com/1981/01/05/world/federal-reserve-sees-little-growth-in-81-with-continued-high-rates.html [Accessed July 07 2021].

THOMPSON, J. M. T. & SIEBER, J. 2011. Climate tipping as a noisy bifurcation: a predictive technique. *IMA Journal of Applied Mathematics,* 76**,** 27-46.

TIAN, S., YU, Y. & GUO, H. 2015. Variable selection and corporate bankruptcy forecasts. *Journal of Banking and Finance,* 52**,** 89-100.

TODA, H. Y. & YAMAMOTO, T. 1995. Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics,* 66**,** 225-250.

VERAART, A. J., FAASSEN, E. J., DAKOS, V., VAN NES, E. H., LÜRLING, M. & SCHEFFER, M. 2012. Recovery rates reflect distance to a tipping point in a living system. *Nature,* 481**,** 357-359.

XIAO, H., ZHU, S., & WANG, L. 2009. Will the financial crisis become the turning point for China's auto industry? A dynamic computable general equilibrium analysis with imperfect competition. *Journal of Chinese Economic and Business Studies*, 7(4), 477-497.

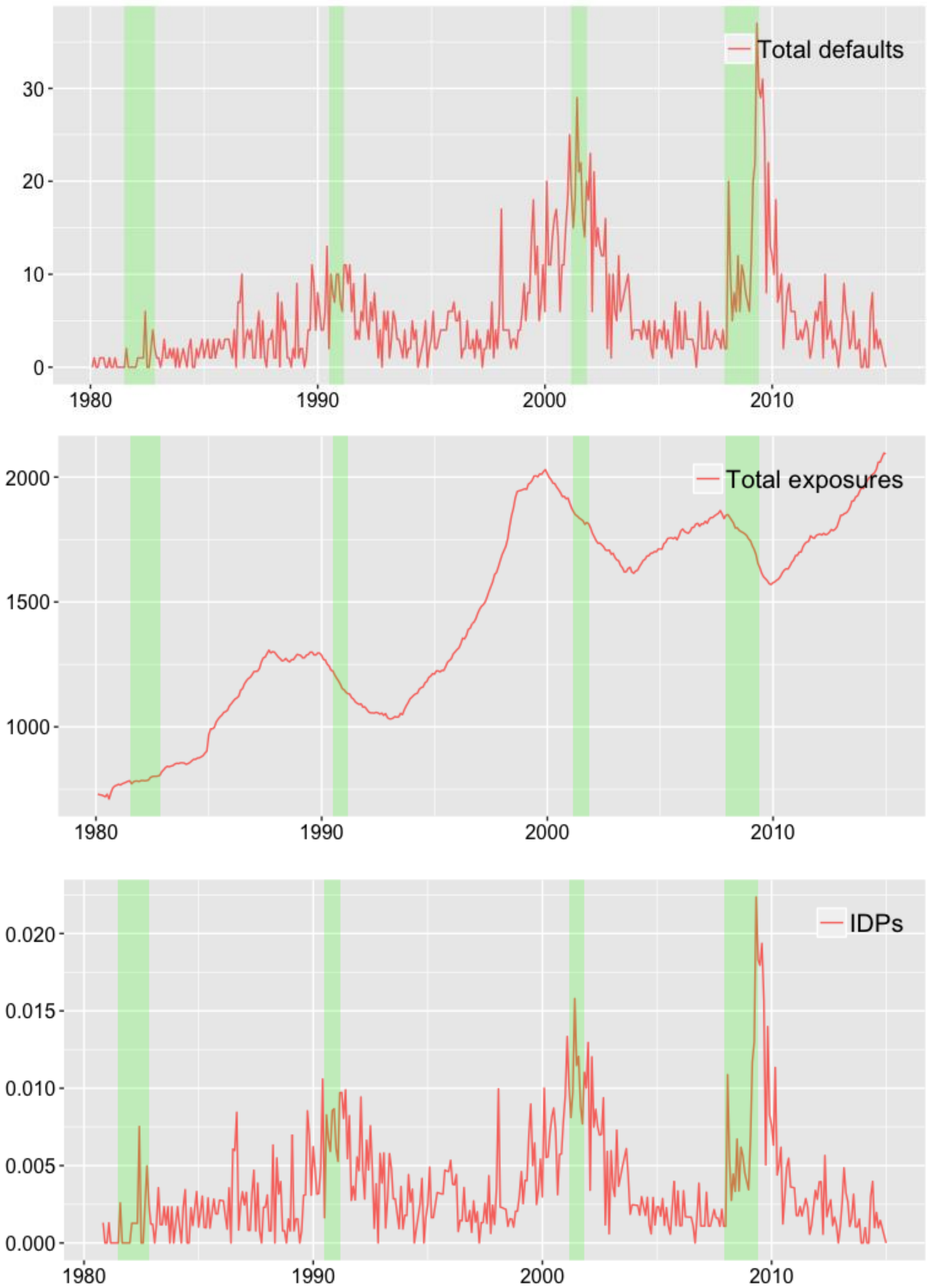
XING, K. & YANG, X. 2019. How to detect crashes before they burst: Evidence from Chinese stock market. *Physica A: Statistical Mechanics and its Applications,* 528**,** 121392.

XING, K. & YANG, X. 2020. Predicting default rates by capturing critical transitions in the macroeconomic system. *Finance Research Letters,* 32**,** 101107.

ZARNOWITZ, V. 1992. Major macroeconomic variables and leading indexes. *Business Cycles: Theory, History, Indicators, and Forecasting.* University of Chicago Press.

ZMIJEWSKI, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research,* 22**,** 59-82.

Figure 1 Time series of monthly default data in US industrial firms from January 1980 to December 2014.



This figure shows time series of total defaults, total exposures, and IDPs in US industrial firms from January 1980 to December 2014. IDPs are fractions in which the numerator represents the number of issuers that defaulted in a particular time period in the first graph and the denominator represents the number of issuers that could have defaulted in that time period in the second graph. The formula of IDP in the third graph is shown below: $IDPs = d_t^{IDP}/n_t^{IDP}$, where $n_t^{IDP} = n_{t-1}^{IDP} - d_{t-1}^{IDP} - w_t^{IDP}$. The numerators d_t^{IDP} and d_{t-1}^{IDP} are the numbers of issuers defaulting at period t and $t-1$. The denominators n_t^{IDP} and n_{t-1}^{IDP} are the numbers of issuers that potentially could have defaulted at date t and $t-1$. w_t^{IDP} denotes the number of credits which withdraw between periods t and $t+1$ (Acharya and Carpenter, 2002, Gansecki, 2010, Moody, 2015).

Table 1 Data description for macro factors

Summary listing	Factors: Full name	Abbreviation	Cyclical factor	Relation with business cycle	8 Economic groups	Full name	Abbreviation	Cyclical factor	Relation with business cycle	8 Economic groups	Total
	Macro-level conditions (85 factors)										
Bank lending conditions	Loans and Leases in Bank Credit, All Commercial Banks	LOANS	Lagging factor	Procyclical	MC	Household debt/income-ratio	TDSP	Lagging factor	Countercyclical	MC	11
	Real Estate Loans, All Commercial Banks,1943	REALLN	Lagging factor	Procyclical	MC	Household obligations/income	FODSP	Lagging factor	Countercyclical	MC	
	Total Consumer Credit Owned and Securitized, Outstanding	TOTALSL	Lagging factor	Procyclical	MC	Interbank Loans, All Commercial Banks	IBLACBM027SBOG	Lagging factor	Procyclical	MC	
	Commercial and Industrial Loans, All Commercial Banks	BUSLOANS	Lagging factor	Procyclical	MC	Borrowings, All Commercial Banks	BOWACBM027SBOG	Lagging factor	Procyclical	MC	
	Consumer Loans at All Commercial Banks	CONSUMER	Lagging factor	Procyclical	MC	Required Reserves of Depository Institutions	REQRESNS	Lagging factor	Countercyclical	MC	
	Federal Debt: Total Public Debt	GFDEBTN	Lagging factor	Countercyclical	MC						MC
General macro indicators	Economic activity index	USPHCI	Coincident factor	Procyclical	PICT	Real Manufacturing and Trade Industries Sales	CMRMTSPL	Coincident factor	Procyclical	PICT	36
	Industrial Production Index	INDPRO	Coincident factor	Procyclical	PICT	Smoothed recession probabilities	RECPROUSM156N	Coincident factor	Countercyclical	NA	
	Industrial Production: Mining: Drilling oil and gas wells	IPN213111S	Coincident factor	Procyclical	PICT	Uni Michigan consumer sentiment	UMCSENT	Leading factor	Procyclical	NA	
	Industrial Production: Manufacturing (SIC)	IPMANSICS	Coincident factor	Procyclical	PICT	Real final sales of domestic product	A190RL1Q225SBEA	Leading factor	Procyclical	PICT	
	Industrial Production: Mining	IPMINE	Coincident factor	Procyclical	PICT	Final Sales to Domestic Purchasers	FSDP	Leading factor	Procyclical	PICT	
	Industrial Production: Electric and Gas Utilities	IPUTIL	Coincident factor	Procyclical	PICT	Expenditure durable goods	PCEDG	Leading factor	Procyclical	PICT	
	Industrial Production: Materials	IPMAT	Coincident factor	Procyclical	PICT	New One Family Houses Sold	HSN1F	Leading factor	Procyclical	FCI	
	Personal income	PI	Coincident factor	Procyclical	PICT	Capacity Utilization: Manufacturing (NAICS)	MCUMFN	Leading factor	Procyclical	PICT	
	Real disposable personal income	DSPIC96	Coincident factor	Procyclical	PICT	Capacity Utilization: Total Industry	TCU	Leading factor	Procyclical	PICT	
	Personal Consumption Expenditures	PCE	Coincident factor	Procyclical	PICT	Moving 12-Month Total Vehicle Miles Traveled	M12MTVUSM227NFWA	Leading factor	Procyclical	PICT	
	Personal Consumption Expenditures: Chain-type Price Index	PCEPI	Coincident factor	Procyclical	PICT	Light Weight Vehicle Sales: Autos & Light Trucks	ALTSALES	Leading factor	Procyclical	PICT	
	Government expenditure	W068RCQ027SBEA	Coincident factor	Countercyclical	INIV	Housing Starts	HOUST	Leading factor	Procyclical	FCI	
	GDP	GDP	Coincident factor	Procyclical	PICT	Building Permits	PERMIT	Leading factor	Procyclical	FCI	
	Gross private domestic investment	GPDI	Coincident factor	Procyclical	FCI	ISM Manufacturing: New Orders Index	NAPMNOI	Leading factor	Procyclical	FCI	
	Private Nonresidential Fixed Investment	PNFI	Coincident factor	Procyclical	FCI	ISM Manufacturing: Inventories Index	NAPMII	Leading factor	Procyclical	INIV	
	Change in private inventories	CBI	Leading factor	Procyclical	INIV	ISM Manufacturing: Supplier Deliveries Index	NAPMSDI	Leading factor	Procyclical	PICT	
	Private Residential Fixed Investment	PRFI	Coincident factor	Procyclical	FCI	ISM manufacturing index	NAPM	Leading factor	Procyclical	PICT	
	Gross National Product	A001RP1Q027SBEA	Coincident factor	Procyclical	PICT	The months' supply is the ratio of houses for sale to houses sold.	MSACSR	Leading factor	Countercyclical	PICT	
Labour market conditions	Initial Claims	ICSA	Leading factor	Countercyclical	EU	Civilian Employment	CE16OV	Coincident factor	Procyclical	EU	13
	Weekly Hours Worked: Manufacturing for the United States	HOHWMN02USM065S	Leading factor	Procyclical	EU	All Employees: Total Nonfarm Payrolls	PAYEMS	Coincident factor	Procyclical	EU	
	Employment Level: Part-Time for Economic Reasons, Slack Work or Business Conditions, Nonagricultural Industries	LNS12032198	Leading factor	Procyclical	EU	All Employees: Manufacturing	MANEMP	Coincident factor	Procyclical	EU	
	ISM Manufacturing: Employment Index	NAPMEI	Leading factor	Procyclical	EU	Average (Mean) Duration of Unemployment	UEMPMEAN	Lagging factor	Countercyclical	EU	
	Civilian Unemployment Rate	UNRATE	Coincident factor	Countercyclical	EU	Of Total Unemployed, Percent Unemployed 27 Weeks and over	LNU03025703	Lagging factor	Countercyclical	EU	
	Labor Market Conditions Index	FRBLMCI	Coincident factor	Procyclical	EU	Number of Civilians Unemployed for 15 Weeks & Over	UEMP15OV	Lagging factor	Countercyclical	EU	
	Civilian Employment-Population Ratio	EMRATIO	Coincident factor	Procyclical	EU			Coincident factor	Procyclical	EU	
Monetary policy indicators	Gross Saving	GSAVE	Coincident factor	Procyclical	MC	Consumer Price Index for All Urban Consumers: Housing	CPIHOSSL	Lagging factor	Procyclical	PCP	15
	Gross Private saving	GPSAVE	Coincident factor	Procyclical	MC	Consumer Price Index for All Urban Consumers: All Items	CPIAUCSL	Lagging factor	Procyclical	PCP	
	Personal Saving	PMSAVE	Coincident factor	Procyclical	MC	Monetary Base	BOGMBASE	Leading factor	Procyclical	MC	
	GDP deflator, implicit	GDPDEF	Coincident factor	Procyclical	PCP	M1 Money Stock	M1SL	Leading factor	Procyclical	MC	
	University of Michigan Inflation Expectation	MICH	Leading factor	Procyclical	PCP	M2 Money Stock	M2SL	Leading factor	Procyclical	MC	
	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy	CPILFESL	Lagging factor	Procyclical	PCP	M3 for the United States	MABMM301USM189S	Leading factor	Procyclical	MC	
	Consumer Price Index for All Urban Consumers: Energy	CPIENGSL	Lagging factor	Procyclical	PCP	MZM Money Stock	MZMSL	Leading factor	Procyclical	MC	
	Consumer Price Index for All Urban Consumers: Transportation	CPITRNSL	Lagging factor	Procyclical	PCP						

Liquidity from non-banks	Retail Money Funds	RMFSL	Leading factor	Procyclical	FCI	Institutional Money Funds	IMFSL	Leading factor	Procyclical	FCI	2
Firm profitability	Corporate Profits After Tax	CP	Lagging factor	Procyclical	PCP	Corporate net cash flow	CNCF	Lagging factor	Procyclical	PCP	3
	Net corporate dividends	B056RC1A027NBEA	Lagging factor	Procyclical	PCP						
Terms of trade	Trade Weighted U.S. Dollar Index: Broad	TWEXBMTH	Leading factor	Procyclical	FTP	Trade Weighted U.S. Dollar Index: Major Currencies	TWEXMMTH	Leading factor	Procyclical	FTP	2
Balance of payments	Net Exports of Goods and Services	NETEXP	Leading factor	Countercyclical	FTP	Real Imports of Goods & Services, 3 Decimal	IMPGSC96	Leading factor	Procyclical	FTP	3
	Real Exports of Goods & Services	EXPGSC1	Leading factor	Countercyclical	FTP						
						Micro-level conditions (29 factors)					
Labour cost/wages	Unit labor cost: nonfarm business	ULCNFB	Coincident factor	Procyclical	PCP	Business Sector: Real Output Per Hour of All Persons	OPHPBS	Coincident factor	Procyclical	PCP	5
	Nonfarm Business Sector: Real Compensation Per Hour	COMPRNFB	Coincident factor	Procyclical	PCP	Nonfarm Business Sector: Real Output Per Hour of All Persons	OPHNFB	Coincident factor	Procyclical	PCP	
	Nonfarm Business Sector: Compensation Per Hour	COMPNFB	Coincident factor	Procyclical	PCP						
Cost of capital	Effective federal funds rate	FEDFUNDS	Lagging factor	Countercyclical	MC	10-Year Treasury Constant Maturity Rate	GS10	Lagging factor	Countercyclical	MC	14
	30 year mortgage rate	MORTG	Lagging factor	Countercyclical	MC	Treasury bond yield, 10 years(Baa)	BAA10YM	Lagging factor	Countercyclical	MC	
	AAA corporate bond yield	AAA	Lagging factor	Countercyclical	MC	3-Month Treasury Bil	TB3MS	Lagging factor	Countercyclical	MC	
	BAA corporate bond yield	BAA	Lagging factor	Countercyclical	MC	10-Year Treasury Constant Maturity Minus Federal Funds Rate	T10YFF	Leading factor	Countercyclical	MC	
	Treasury bond yield, 10 years(Aaa)	AAA10YM	Lagging factor	Countercyclical	MC	Interest spead: Difference between 10-year Treasury constant maturity rate and 1-year Treasury constant maturity rate	SPREAD.GS	Leading factor	Countercyclical	MC	
	Bank prime loan rate	MPRIME	Lagging factor	Countercyclical	MC	Interest spead: Difference between 1-year BAA yield and 1-year AAA yield	SPREAD.MOODY.1	Leading factor	Countercyclical	MC	
	1-Year Treasury Constant Maturity Rate	GS1	Lagging factor	Countercyclical	MC	Interest spead: Difference between 10-year BAA yield 10-year and AAA yield	SPREAD.MOODY.2	Leading factor	Countercyclical	MC	
Cost of resources	PPI all commodities	PPIACO	Lagging factor	Procyclical	PCP	PPI industrial commodities	PPIIDC	Lagging factor	Procyclical	PCP	6
	PPI interm. energy goods	PPIIEG	Lagging factor	Procyclical	PCP	PPI crude energy materials	PPICEM	Lagging factor	Procyclical	PCP	
	PPI finished goods	PPIFGS	Lagging factor	Procyclical	PCP	PPI intermediate materials	PPIITM	Lagging factor	Procyclical	PCP	
Equity indexes and respective volatilities	SP500 index	SP500	Leading factor	Procyclical	-	Russell 2000 index	RU2000	Leading factor	Procyclical	-	4
	NASDAQ composite index	NASDAQ	Leading factor	Procyclical	-	Wilshire 5000 Total Market Full Cap Index	WILL5000	Leading factor	Procyclical	-	
					Total						114

Note: In the column named "8 Economic groups", we follow the method proposed by Shiskin and Moore (1967) to identify macro factors into 8 groups, including (1) Money and credit is short for MC; (2) Production, income, consumption, and trade is short for PICT; (3) Federal government activities is short for FGA; (4) Inventories and inventory investment is short for INIV; (5) Fixed capital investment is short for FCI; (6) Employment and unemployment is short for EU; (7) Foreign trade and payments is short for FTP; (8) Prices, cost, and profits is short for PCP.. Since there are no factors in economic activity in other economy, then this study only classifies all the macro factors into 8 economic groups. The classification of "cyclical factors" is mainly based on Shiskin and Moore (1967), and *Business Cycle Indicators Handbook* published by the Conference Board also gives a detailed results of classification of hundreds of macroeconomic factors. We assign each macro factor into 1 of 3 groups including leading, coincident, and lagging factors according to the timing of their movements. For instance, leading factors tend to shift direction in advance of the business cycle. The detailed description of each factor available upon request.

Table 2 Classifications of macroeconomic conditions based on different combinations of macro factors

Combinations of macro factors (Number of factors)	Types of macroeconomic conditions
All the factors (114)	Augmented macroeconomic conditions
Factors in 6 economic groups (106)	Generalized macroeconomic conditions
Leading factors (42) Procyclical factors (86)	Incomplete macroeconomic conditions
Effective factors (29)	Specific macroeconomic conditions

Note: Both augmented macroeconomic conditions and generalized macroeconomic conditions are able to mirror comparatively complete macroeconomic conditions.

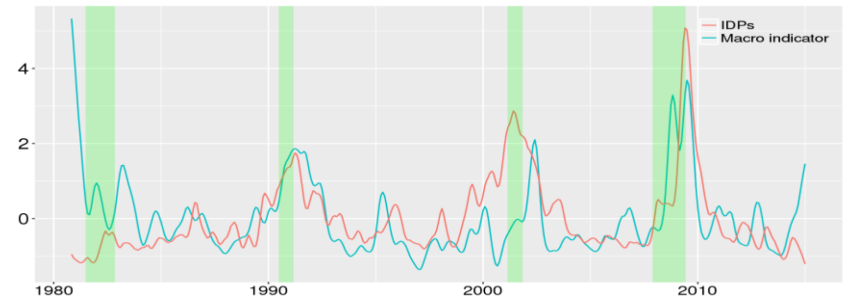
Table 3 Summary statistics of IDP, defaults, and various indicators from September 1980 to December 2014

Panel A Before filtering								
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
IDP	412	0.004	0.003	0	0.001	0.003	0.005	0.022
Defaults	412	5.432	5.685	0	2	4	7	37
Indicator_all	412	0.158	0.028	0.115	0.14	0.149	0.167	0.289
Indicator_leading	412	0.144	0.029	0.102	0.128	0.138	0.151	0.308
Indicator_procyclical	412	0.162	0.031	0.113	0.141	0.153	0.172	0.305
Indicator_6_economic	412	0.156	0.028	0.113	0.138	0.148	0.165	0.287
Indicator_effective	412	0.167	0.041	0.088	0.139	0.16	0.184	0.349
Panel B After filtering								
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
IDP	412	0.004	0.003	0.0001	0.002	0.002	0.005	0.018
Indicator_all	412	0.158	0.025	0.124	0.141	0.15	0.164	0.288
Indicator_leading	412	0.144	0.025	0.108	0.131	0.138	0.148	0.302
Indicator_procyclical	412	0.162	0.028	0.124	0.143	0.153	0.168	0.29
Indicator_6_economic	412	0.156	0.025	0.123	0.139	0.149	0.163	0.291
Indicator_effective	412	0.167	0.036	0.11	0.142	0.161	0.182	0.309

Note: Filtering means that we use HP filtering method to filter the time series in order to remove the cyclical components.

Figure 2 Dynamics of each macro indicator and IDPs from Oct 1980 to Dec 2014 in US industrial firms and breakpoints in the time series of log-transformed IDPs

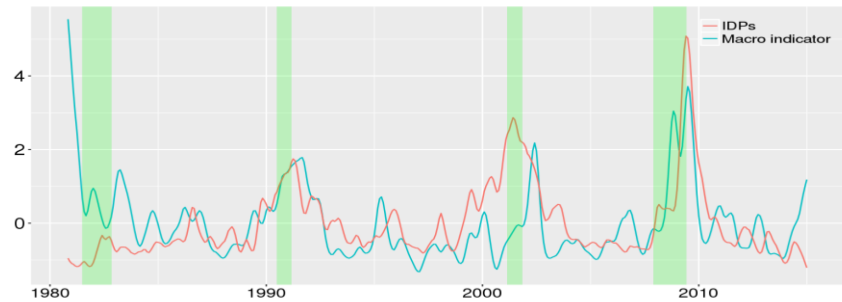
Panel A: All the factors



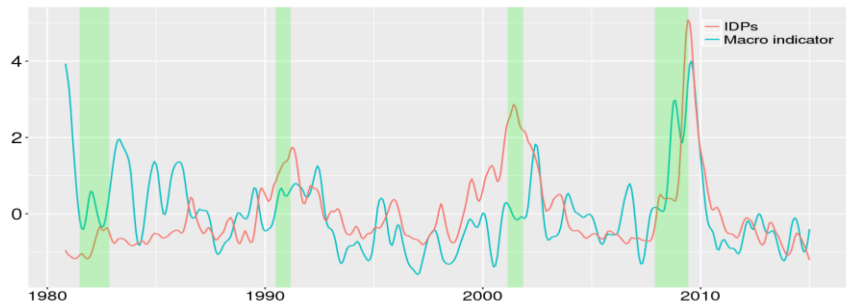
Panel B: Leading factors

Panel C: Procyclical factors

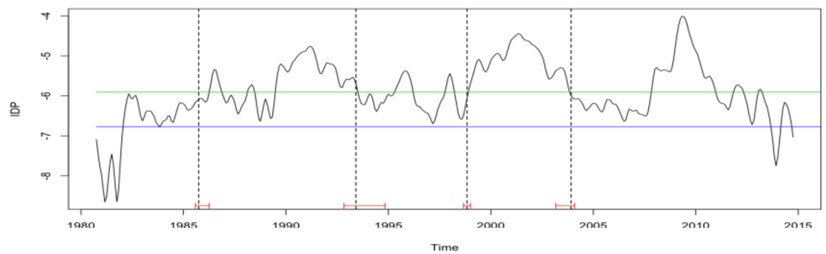
Panel D: 6 Economic groups



Panel E: Effective factors



Panel F: Breakpoints in series of log-transformed IDP



Panel F shows the breakpoints in the time series of log-transformed IDP. This paper uses the approach proposed by Bai and Perron (2003) to determine 4 breakpoints, which are occurred in October 1985, June 1993, November 1998, December 2003. The best combination of breakpoints is identified by BIC and they are highly significant. Three points occurred after three economic recessions. The third point occurred after the Asian financial crisis in 1997.

Table 4 Confirmatory Analysis

Interval	IDPs	Indicator (All factors)		Indicator (Leading factors)		Indicator (Procyclical factors)		Indicator (6 Economic groups)		Indicator (Effective factors)	
	I(m)	I(m)	Types of tests	I(m)	Types of tests	I(m)	Types of tests	I(m)	Types of tests	I(m)	Types of tests
1980.09-2014.12	I(1)	I(0)		I(0)		I(0)		I(1)	CT	I(1)	CT
1980.09-1985.09	I(1)	I(4)		I(1)	CT	I(1)	CT	I(7)		I(0)	
1985.10-1993.06	I(1)	I(1)	CT	I(1)	CT	I(1)	CT	I(1)	CT	I(0)	
1993.07-1998.11	I(0)	I(0)	GCT	I(1)		I(0)	GCT	I(1)		I(1)	
1998.12-2003.12	**I(1)**	I(0)		I(0)		I(0)		I(0)		I(1)	CT
2004.01-2014.12	I(1)	I(0)		I(0)		I(0)		I(0)		I(1)	CT

Note: I(m) means that the variable converts to stationarity after being integrated of the order m. The bold type I(1) means that after using various order for integration in the variables, these variables still do not pass the stationary test based on KPSS. However, they pass the ADF test after first order integration. Although this study uses KPSS test as the standard for deciding the order of m, for this case, the ADF test is used. Therefore, the cointegration test is used for cross checking whether they are cointegrated with the other variables.

Two abbreviations in the column titled Types of tests are CT and GCT, which represent cointegration test and Granger causality test, respectively. If both two variables are stationary, then Granger causality is used for testing the causality relationship between two variables (Lütkepohl, 2005). For the other cases, Toda and Yamamoto causality test is used for checking causality in the nonstationary variables (Pfaff, 2008).

Table 5 Johansen's cointegration tests between each macro indicator and IDPs

This table provides the cointegration results after using Johansen's test in 6 samples including the full sample, sample from September 1980 to September 1985, sample from October 1985 to June 1993, sample from July 1993 to November 1998, sample from December 1998 to December 2003, and sample from January 2004 to December 2014. Panel A reports the results from using eigenvalue test between each indicator and IDP in different samples. Panel B reports the results from using trace test in different samples. $r = 0$, there is no cointegrated vectors. $r <= 1$, there is 1 integrated vector. FH means that the test fails to reject $r = 0$ or $r <= 1$. Significant at the ***1% level, the ** 5% level and the * 10% level.

Panel A Eigenvalue Test														
		Values of test statistic/Decision										Critical values of test		
Interval		Indicator (All factors)		Indicator (Leading factors)		Indicator (Procyclical factors)		Indicator (6 Economic groups)		Indicator (Effective factors)		10%	5%	1%
1980.09-2014.12	$r <= 1$							**0.91**	**FH**	**1.83**	**FH**	7.52	9.24	12.97
	$r = 0$							**24.13**	**RH**	**20.08**	**RH**	13.75	15.67	20.2
1980.09-1985.09	$r <= 1$			10.36	RH	9.25	RH					7.52	9.24	12.97
	$r = 0$			23.74	RH	12.84	FH					13.75	15.67	20.2
1985.10-1993.06	$r <= 1$	2.14	FH	1.99	FH	1.94	FH	2.59	FH			7.52	9.24	12.97
	$r = 0$	9.71	FH	10.46	FH	9.76	FH	9.15	FH			13.75	15.67	20.2
1993.07-1998.11	$r <= 1$											7.52	9.24	12.97
	$r = 0$											13.75	15.67	20.2
1998.12-2003.12	$r <= 1$									**0.63**	**FH**	7.52	9.24	12.97
	$r = 0$									**18.96**	**RH**	13.75	15.67	20.2
2004.01-2014.12	$r <= 1$									2.87	FH	7.52	9.24	12.97
	$r = 0$									8.89	FH	13.75	15.67	20.2
Panel B Trace Test														
		Values of test statistic										Critical values of test		
Interval		Indicator (All factors)		Indicator (Leading factors)		Indicator (Procyclical factors)		Indicator (6 Economic groups)		Indicator (Effective factors)		10%	5%	1%
1980.09-2014.12	$r <= 1$							**0.91**	**FH**	**1.83**	**FH**	7.52	9.24	12.97
	$r = 0$							**25.04**	**RH**	**21.91**	**RH**	17.85	19.96	24.6
1980.09-1985.09	$r <= 1$			10.36	RH	9.25	RH					7.52	9.24	12.97
	$r = 0$			34.11	RH	22.09	RH					17.85	19.96	24.6
1985.10-1993.06	$r <= 1$	2.14	FH	1.99	FH	1.94	FH	2.59	FH			7.52	9.24	12.97
	$r = 0$	11.84	FH	12.45	FH	11.7	FH	11.74	FH			17.85	19.96	24.6
1993.07-1998.11	$r <= 1$											7.52	9.24	12.97
	$r = 0$											17.85	19.96	24.6
1998.12-2003.12	$r <= 1$									**0.63**	**FH**	7.52	9.24	12.97
	$r = 0$									**19.58**	**RH**	17.85	19.96	24.6
2004.01-2014.12	$r <= 1$									2.87	FH	7.52	9.24	12.97
	$r = 0$									11.76	FH	17.85	19.96	24.6

Table 6 Causality analysis

This table shows results of investigation of the causal relationship between IDPs and macro indicators in 6 samples including the full sample, sample from September 1980 to September 1985, sample from October 1985 to June 1993, sample from July 1993 to November 1998, sample from December 1998 to December 2003, sample from January 2004 to December 2014. The results of max order integration ($\boldsymbol{m}$) is from Table 4 and the results of lag in VAR ($\boldsymbol{p}$) is from **Table A2 in the appendix**. Following the results from **Table A1 in the appendix**, all the indicators except the indicator constructed by leading factors in the fourth interval and IDP are stationary in the levels I(0). Then their relationships are tested by Granger causality. For the other cases, this study uses TY causality test proposed by Toda and Yamamoto (1995) for investigating the relationship between each indicator and IDPs. Shadow area means that there is cointegration between IDPs and macro indicator. Significant at the ***1% level, the ** 5% level and the * 10% level.

Panel A Macro indicator constructed by all the factors

	H_0: Indicator does not Granger-cause IDPs		H_0: IDPs does not Granger-cause Indicator		Max order integration (m)	Lag in VAR (p)
Interval	Test Statistic	P-value	Test Statistic	P-value		
1980.09-2014.12	**23.9**	**0.00055*** **	**11.5**	**0.074***	1	6
1980.09-1985.09	5.7	0.22	**14**	**0.0071*** **	4	4
1985.10-1993.06	3	0.56	**10.4**	**0.034*** *	1	4
1993.07-1998.11	**3.4067**	**0.0151*** *	1.1143	0.3598	1	4
1998.12-2003.12	2.9	0.58	1.7	0.79	1	4
2004.01-2014.12	**21.3**	**0.00027*** **	6.4	0.17	1	4

Panel B Macro indicator constructed by leading factors

	H_0: Indicator does not Granger-cause IDPs		H_0: IDPs does not Granger-cause Indicator		Max order integration (m)	Lag in VAR (p)
Interval	Test Statistic	P-value	Test Statistic	P-value		
1980.09-2014.12	**15**	**0.0047*** **	0.74	0.95	1	4
1980.09-1985.09	**17.6**	**0.0015*** **	2.1	0.71	1	4
1985.10-1993.06	0.54	0.97	**10.6**	**0.031*** *	1	4
1993.07-1998.11	3.9	0.42	2.3	0.67	1	4
1998.12-2003.12	2.7	0.44	2.3	0.52	1	3
2004.01-2014.12	**17.2**	**0.0018*** **	4.4	0.35	1	4

Panel C Macro indicator constructed by procyclical factors

	H_0: Indicator does not Granger-cause IDPs		H_0: IDPs does not Granger-cause Indicator		Max order integration (m)	Lag in VAR (p)
Interval	Test Statistic	P-value	Test Statistic	P-value		
1980.09-2014.12	**17.7**	**0.0014*** **	5.4	0.25	1	4
1980.09-1985.09	**9.8**	**0.044*** *	**10.1**	**0.039*** *	1	4
1985.10-1993.06	1.9	0.76	5.9	0.21	1	4
1993.07-1998.11	2.022	0.1215	**5.68**	**0.0018*** **	1	3
1998.12-2003.12	2.6	0.63	1	0.91	1	4
2004.01-2014.12	**14.1**	**0.0069*** **	4	0.4	1	4

Panel D Macro indicator constructed by the factors in the 6 economic groups

	H_0: Indicator does not Granger-cause IDPs		H_0: IDPs does not Granger-cause Indicator		Max order integration (m)	Lag in VAR (p)
Interval	Test Statistic	P-value	Test Statistic	P-value		
1980.09-2014.12	**24.9**	**0.000053*** **	5.3	0.26	1	4
1980.09-1985.09	2.3	0.67	**11**	**0.027*** *	7	4
1985.10-1993.06	1.9	0.75	**10**	**0.04*** *	1	4
1993.07-1998.11	2.2	0.7	1.9	0.76	1	4
1998.12-2003.12	3.9	0.42	1.2	0.88	1	4
2004.01-2014.12	**20.8**	**0.00034*** **	6.7	0.15	1	4

Panel E Macro indicator constructed by the effective factors extracted by Lasso

	H_0: Indicator does not Granger-cause IDPs		H_0: IDPs does not Granger-cause Indicator		Max order integration (m)	Lag in VAR (p)
Interval	Test Statistic	P-value	Test Statistic	P-value		
1980.09-2014.12	**15.2**	**0.0094*** **	8.6	0.12	1	5
1980.09-1985.09	**10.4**	**0.034*** *	3.5	0.48	1	4
1985.10-1993.06	1.9	0.76	3.1	0.54	1	4
1993.07-1998.11	**3.9192**	**0.0074*** **	1.1132	0.3603	1	4
1998.12-2003.12	7.1	0.13	2.2	0.7	1	4
2004.01-2014.12	**16.3**	**0.0027*** **	9.4	**0.051***	1	4

Figure 3 Dynamic of macro indicator and FIDPs after removing the monthly defaults (<= 5) (with no lag for each variable)

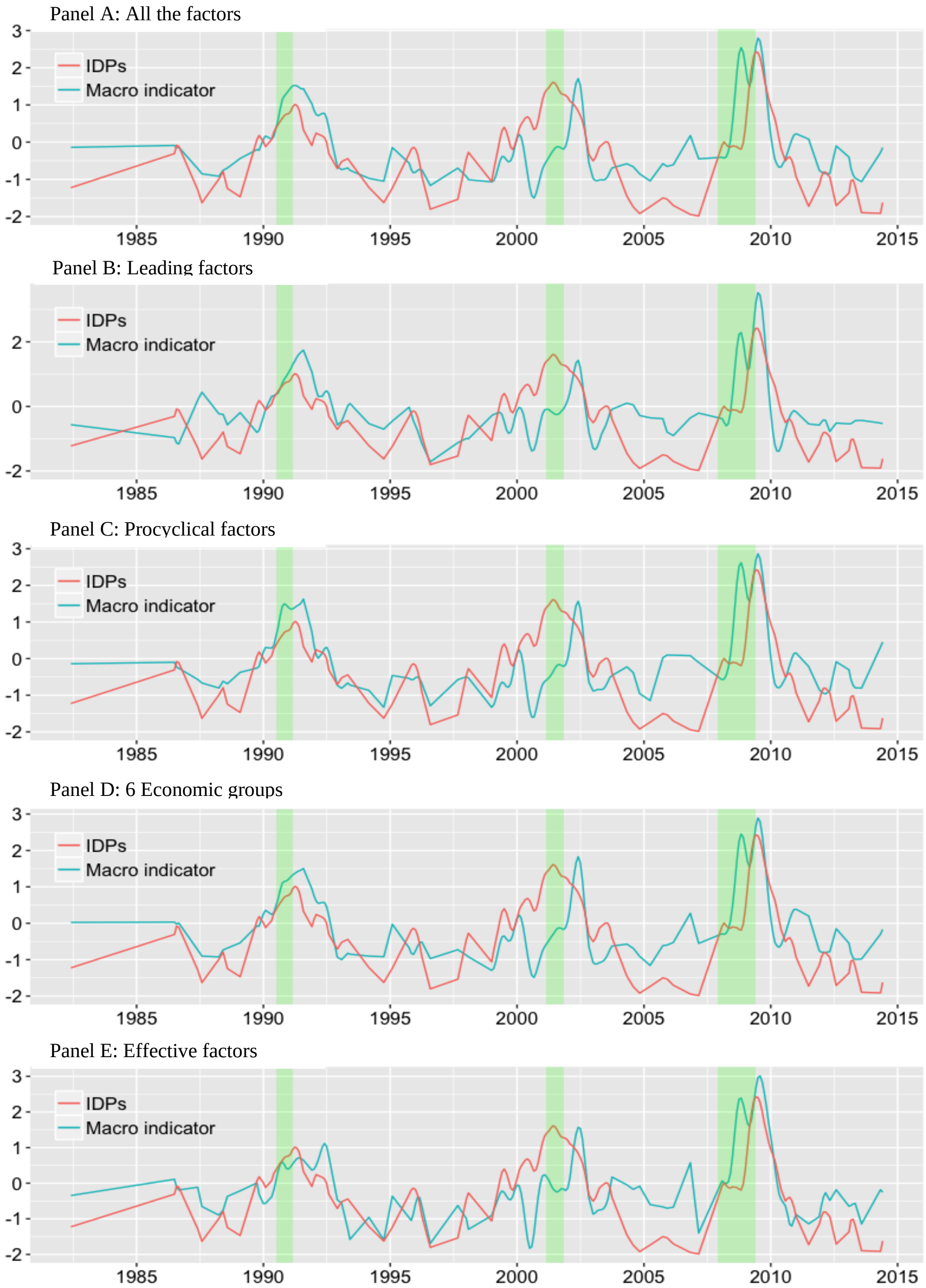


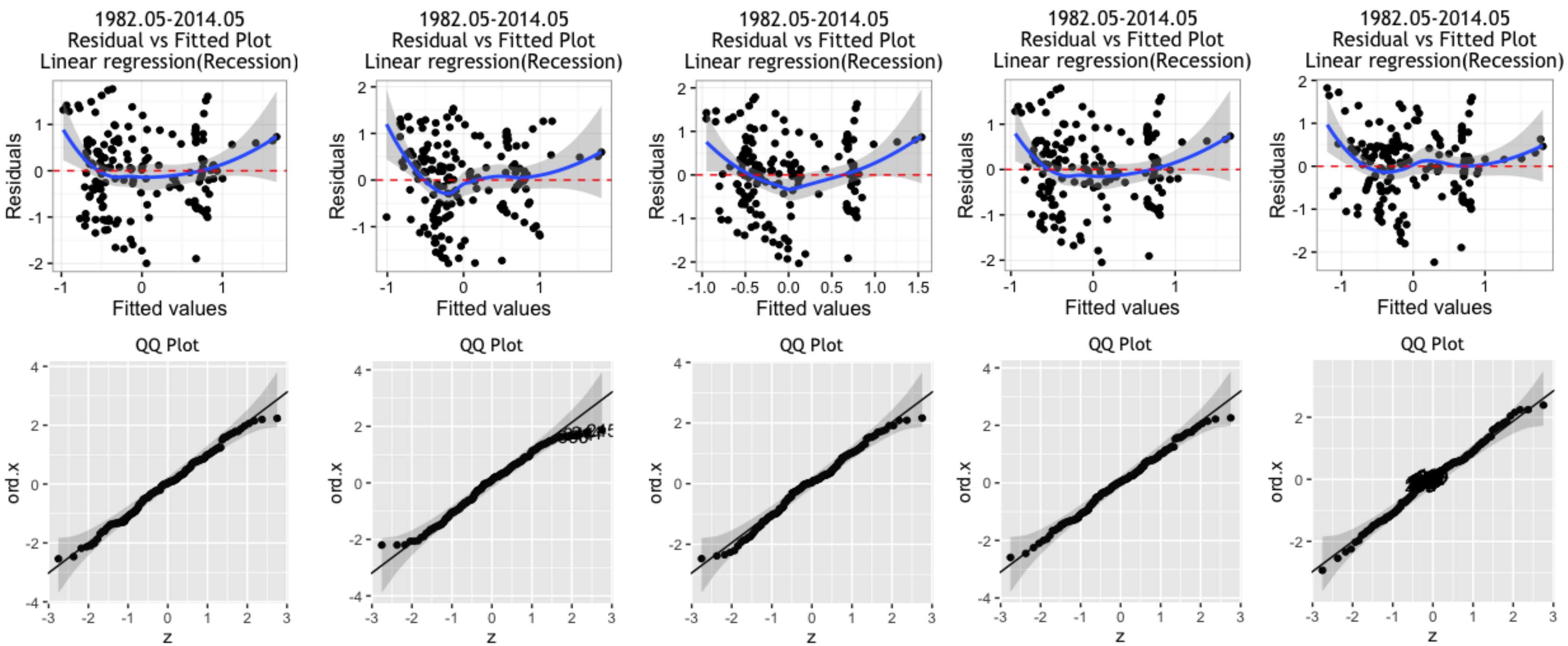
Table 7 Regression results without any lags

Panel A reports the regression results. The regression is defined as $Y_t = \beta_0 + \beta_1 X_t + \beta_2 D_t + \beta_3 X_t D_t + \varepsilon$, where Y_t is the dependent variable FIDPs at date t, and X_t is independent variable (5 macro indicators) at date t. Both Y and X are transformed to logarithmic ratios for stabilization. D_t is the dummy variable, 1 means recession; 0 means no recession. The macro indicators in the column from (1) to (5) are constructed by five groups of factors, which are all the factors, leading factors, procyclical factors, factors in the 6 economic groups, and effective factors, respectively. Panel B reports the mean default risk (FIDPs) function in nonrecession period. The function is written as $E(Y_t|D_t = 0, X_t) = \beta_0 + \beta_1 X_t$. Panel C reports the mean default risk (FIDPs) function in recession period. The function is defined as $E(Y_t|D_t = 1, X_t) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)X_t$, where $\beta_0 + \beta_2$ is the new constant and $\beta_1 + \beta_3$ is the new coefficient for indicator, X_t. Both FIDPs and each indicator are standardised. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. Standard errors are reported in parentheses underneath the coefficients.

Panel A Regression results					
	Dependent variable: FIDPs				
	(1) All the factors	(2) Leading factors	(3) Procyclical factors	(4) 6 Economic groups	(5) Effective factors
Indicator	0.616***	0.537***	0.558***	0.597***	0.620***
	(0.079)	(0.078)	(0.086)	(0.079)	(0.074)
Recession	0.732***	0.689***	0.752***	0.737***	0.751***
	(0.193)	(0.187)	(0.200)	(0.194)	(0.187)
Indicator *Recession	-0.563***	-0.363**	-0.522***	-0.541***	-0.570***
	(0.157)	(0.160)	(0.158)	(0.161)	(0.161)
Constant	-0.051	-0.087	-0.059	-0.058	-0.064
	(0.071)	(0.072)	(0.075)	(0.071)	(0.069)
Observations	169	169	169	169	169
R2	0.365	0.331	0.31	0.358	0.394
Adjusted R2	0.354	0.319	0.297	0.346	0.383
Residual Std. Error	0.804 (df = 165)	0.825 (df = 165)	0.838 (df = 165)	0.808 (df = 165)	0.786 (df = 165)
F Statistics	31.658*** (df = 3; 165)	27.226*** (df = 3; 165)	24.674*** (df = 3; 165)	30.679*** (df = 3; 165)	35.725*** (df = 3; 165)
Panel B Mean default risk (FIDPs) function for non-recession					
Indicator	0.616	0.537	0.558	0.597	0.62
Constant	-0.051	-0.087	-0.059	-0.058	-0.064
Panel C Mean default risk (FIDPs) function for recession					
Indicator	0.053	0.174	0.036	0.056	0.05
Constant	0.681	0.602	0.693	0.679	0.687

Figure 4 Model diagnostics for the models in Table 7

Panel A1: All the factors | Panel A2: Leading factors | Panel A3: Procyclical factors | Panel A4: 6 Economic groups | Panel A5: Effective factors



Panel B: Results of heteroscedasticity tests

Types of indicator	Statistic	P-value
(1) All the factors	1.8876	0.1695
(2) Leading factors	1.3029	0.2537
(3) Procyclical factors	0.5499	0.4584
(4) 6 Economic groups	1.6843	0.1944
(5) Effective factors	1.5838	0.2082

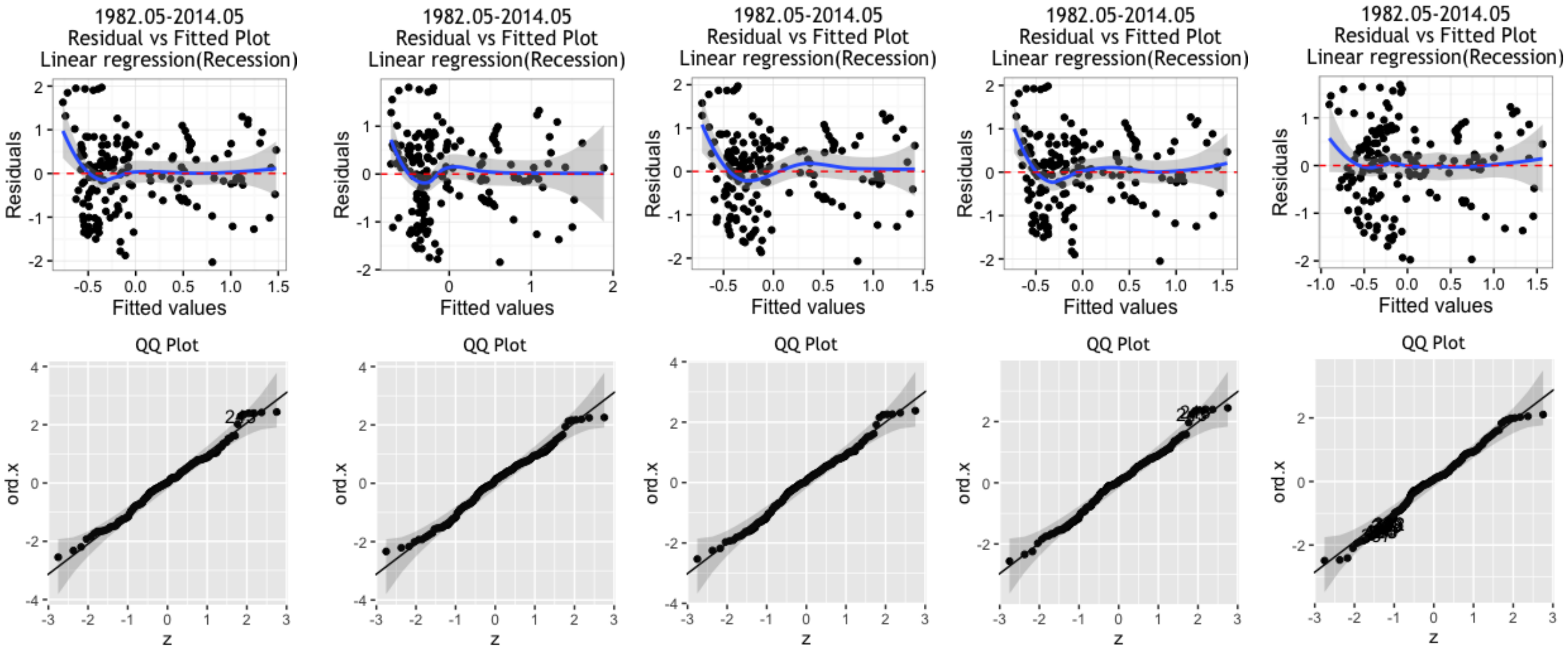
Table 8 Regression results after lagging indicators for 3-month

Panel A reports the regression results. The regression is defined as $Y_t = \beta_0 + \beta_1 X_{t-3} + \beta_2 D_{t-3} + \beta_3 X_{t-3} D_{t-3} + \varepsilon$, where Y_t is the dependent variable FIDP at date $t-3$, and X_{t-3} is independent variable (5 macro indicators) at date $t-3$. Both Y and X are transformed to logarithmic ratios for stabilization. D_t is the dummy variable, 1 means recession; 0 means no recession. The macro indicators in the column from (1) to (5) are constructed by five groups of factors, which are all the factors, leading factors, procyclical factors, factors in the 6 economic groups, and effective factors, respectively. Panel B reports the mean default risk (FIDPs) function in nonrecession period. The function is written as $E(Y_t|D_{t-3} = 0, X_{t-3}) = \beta_0 + \beta_1 X_{t-3}$. Panel C reports the mean default risk (FIDPs) function in recession period. The function is defined as $E(Y_t|D_{t-3} = 1, X_{t-3}) = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_{t-3}$, where $\beta_0 + \beta_2$ is the new constant and $\beta_1 + \beta_3$ is the new coefficient for indicator, X_{t-3}. Both FIDP and each indicator are standardised. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. Standard errors are reported in parentheses underneath the coefficients.

Panel A Regression results					
	Dependent variable: FIDPs				
	(1) All the factors	(2) Leading factors	(3) Procyclical factors	(4) 6 Economic groups	(5) Effective factors
Indicator	0.404***	0.388***	0.354***	0.386***	0.403***
	(0.083)	(0.079)	(0.088)	(0.081)	(0.077)
Recession	0.775***	0.791***	0.808***	0.769***	0.797***
	(0.195	(0.183)	(0.198)	(0.196)	(0.193)
Indicator *Recession	-0.08	-0.006	-0.073	-0.044	-0.06
	(0.158)	(0.156)	(0.158)	(0.162)	0.168)
Constant	-0.146**	-0.163**	-0.154**	-0.152**	-0.156**
	(0.073)	(0.072)	(0.076)	(0.074)	0.072)
Observations	169	169	169	169	169
R2	0.334	0.343	0.302	0.329	0.344
Adjusted R2	0.321	0.331	0.29	0.317	0.332
Residual Std. Error	0.824 (df = 165)	0.818 (df = 165)	0.843 (df = 165)	0.826 (df = 165)	0.818 (df = 165)
F Statistic	27.532*** (df = 3; 165)	28.729*** (df = 3; 165)	23.843*** (df = 3; 165)	26.992*** (df = 3; 165)	28.793*** (df = 3; 165)
Panel B Mean default risk (FIDPs) function for non-recession					
Indicator	0.404	0.388	0.354	0.386	0.403
Constant	-0.146	-0.163	-0.154	-0.152	-0.156
Panel C Mean default risk (FIDPs) function for recession					
Indicator	0.324	0.382	0.281	0.342	0.343
Constant	0.629	0.628	0.654	0.617	0.641

Figure 5 Model diagnostics for the models in Table 8

Panel A1: All the factors | Panel A2: Leading factors | Panel A3: Procyclical factors | Panel A4: 6 Economic groups | Panel A5: Effective factors



Panel B: Results of heteroscedasticity tests

Types of indicator	Statistic	P-value
(1) All the factors	2.7578	0.0968
(2) Leading factors	1.7793	0.1822
(3) Procyclical factors	1.1134	0.2914
(4) 6 Economic groups	2.5978	0.107
(5) Effective factors	0.2486	0.618

Appendix

Table A1 Tests of the unit root hypothesis

	Interval	Test	Statistic	Lags	*p*-value
	Levels				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	0.3539	7	0.99
		KPSS	0.9408	4	0.01**
	1980.09-1985.09	ADF	-2.256	3	0.4713
		KPSS	1.5659	1	0.01**
	1985.10-1993.06	ADF	-1.7906	4	0.6626
		KPSS	1.3428	2	0.01**
	1993.07-1998.11	ADF	-2.8759	3	0.2199
		KPSS	0.2675	1	0.1
	1998.12-2003.12	ADF	-0.7864	3	0.9578
		KPSS	0.7142	1	0.0123**
	2004.01-2014.12	ADF	0.7326	5	0.99
		KPSS	0.7789	2	0.01**
IDPs	Differences				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-4.7317	7	0.01**
		KPSS	0.2249	4	0.1
	1980.09-1985.09	ADF	-5.1475	3	0.01**
		KPSS	0.0757	1	0.1
	1985.10-1993.06	ADF	-4.0731	4	0.01**
		KPSS	0.0737	2	0.1
	1993.07-1998.11	ADF	-3.1668	3	0.1022
		KPSS	0.0764	1	0.1
	1998.12-2003.12	ADF	-3.7363	3	0.0293**
		KPSS	0.9202	1	0.01**
	2004.01-2014.12	ADF	-0.2466	5	0.99
		KPSS	0.3476	2	0.0998*
	Levels				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-2.9714	7	0.167
		KPSS	0.3862	4	0.0831*
	1980.09-1985.09	ADF	-3.7092	3	0.0315**
		KPSS	1.559	1	0.01**
	1985.10-1993.06	ADF	-1.1688	4	0.9073
		KPSS	0.8577	2	0.01**
	1993.07-1998.11	ADF	-2.3921	3	0.4161
		KPSS	0.3642	1	0.0926*
	1998.12-2003.12	ADF	-2.7408	3	0.2752
		KPSS	0.2884	1	0.1
	2004.01-2014.12	ADF	-2.1193	5	0.5269
		KPSS	0.4107	2	0.0725*
Indicator_all	Differences: Bold color means that the time series is stable after 4th difference				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-5.9242	7	0.01**
		KPSS	0.2508	4	0.1
	1980.09-1985.09	ADF	-4.5411	3	0.01**
		KPSS	0.4493	1	0.0559*
	1985.10-1993.06	ADF	-2.0728	4	0.5463
		KPSS	0.3585	2	0.095*
	1993.07-1998.11	ADF	-4.0314	3	0.014**
		KPSS	0.1293	1	0.1
	1998.12-2003.12	ADF	-2.8114	3	0.2468
		KPSS	0.1392	1	0.1
	2004.01-2014.12	ADF	-2.4142	5	0.4043
		KPSS	0.1833	2	0.1
	Levels				
Indicator_leading	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-3.7933	7	0.0195**
		KPSS	0.2791	4	0.1

	1980.09-1985.09	ADF	-4.1779	3	0.01**
		KPSS	1.0619	1	0.01**
	1985.10-1993.06	ADF	-1.9183	4	0.61
		KPSS	1.0752	2	0.01**
	1993.07-1998.11	ADF	-2.7255	3	0.2809
		KPSS	0.7892	1	0.01**
	1998.12-2003.12	ADF	-2.591	3	0.3358
		KPSS	0.2491	1	0.1
	2004.01-2014.12	ADF	-2.1678	5	0.5067
		KPSS	0.2484	2	0.1
	Differences				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-6.4471	7	0.01**
		KPSS	0.2058	4	0.1
	1980.09-1985.09	ADF	-4.2964	3	0.01**
		KPSS	0.8115	1	0.01**
	1985.10-1993.06	ADF	-3.6763	4	0.0308**
		KPSS	0.0953	2	0.1
	1993.07-1998.11	ADF	-2.5971	3	0.333
		KPSS	0.3133	1	0.1
	1998.12-2003.12	ADF	-3.8178	3	0.0237**
		KPSS	0.0761	1	0.1
	2004.01-2014.12	ADF	-3.0605	5	0.1356
		KPSS	0.1199	2	0.1
Indicator_procyclical	Levels				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-3.2508	7	0.0794*
		KPSS	0.3525	4	0.0976*
	1980.09-1985.09	ADF	-3.7332	3	0.0295**
		KPSS	1.9034	1	0.01**
	1985.10-1993.06	ADF	-0.7675	4	0.9613
		KPSS	0.6867	2	0.0148**
	1993.07-1998.11	ADF	-2.4886	3	0.3769
		KPSS	0.1822	1	0.1
	1998.12-2003.12	ADF	-2.7776	3	0.2603
		KPSS	0.4063	1	0.0744*
	2004.01-2014.12	ADF	-2.0284	5	0.5647
		KPSS	0.3133	2	0.1
	Differences				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-6.4996	7	0.01**
		KPSS	0.2101	4	0.1
	1980.09-1985.09	ADF	-3.6821	3	0.0339**
		KPSS	0.3493	1	0.099*
	1985.10-1993.06	ADF	-2.7432	4	0.2699
		KPSS	0.2711	2	0.1
	1993.07-1998.11	ADF	-3.3089	3	0.0783*
		KPSS	0.1953	1	0.1
	1998.12-2003.12	ADF	-2.6991	3	0.2922
		KPSS	0.1307	1	0.1
	2004.01-2014.12	ADF	-3.0569	5	0.137
		KPSS	0.1551	2	0.1
Indicator_6_economic.groups	Levels				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-2.9573	7	0.173
		KPSS	0.4679	4	0.0489**
	1980.09-1985.09	ADF	-3.974	3	0.0168**
		KPSS	1.6245	1	0.01**
	1985.10-1993.06	ADF	-1.1944	4	0.9032
		KPSS	0.6223	2	0.0206**
	1993.07-1998.11	ADF	-2.5297	3	0.3603
		KPSS	0.4902	1	0.0439**
	1998.12-2003.12	ADF	-2.8759	3	0.2206
		KPSS	0.2779	1	0.1

	Interval	Test	Statistic	Lags	*p*-value
	2004.01-2014.12	ADF	-2.0494	5	0.5559
		KPSS	0.4379	2	0.0608*
	Differences: Bold color means that the time series is stable after 7th difference				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-6.0482	7	0.01**
		KPSS	0.253	4	0.1
	1980.09-1985.09	ADF	-2.7313	3	0.28
		KPSS	0.4609	1	0.0509*
	1985.10-1993.06	ADF	-2.1385	4	0.5192
		KPSS	0.2951	2	0.1
	1993.07-1998.11	ADF	-4.1732	3	0.01**
		KPSS	0.1549	1	0.1
	1998.12-2003.12	ADF	-2.6952	3	0.2938
		KPSS	0.1441	1	0.1
	2004.01-2014.12	ADF	-2.5251	5	0.3582
		KPSS	0.1806	2	0.1
Indicator_effective	Levels				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-3.472	7	0.0453**
		KPSS	0.5993	4	0.0227**
	1980.09-1985.09	ADF	-3.2172	3	0.0932*
		KPSS	0.2192	1	0.1
	1985.10-1993.06	ADF	-2.336	4	0.4377
		KPSS	0.2381	2	0.1
	1993.07-1998.11	ADF	-3.3067	3	0.0785*
		KPSS	0.4636	1	0.0499**
	1998.12-2003.12	ADF	-3.3271	3	0.0758*
		KPSS	0.4794	1	0.0463**
	2004.01-2014.12	ADF	-1.9064	5	0.6154
		KPSS	0.5077	2	0.0399**
	Differences				
	Interval	Test	Statistic	Lags	*p*-value
	1980.09-2014.12	ADF	-7.1328	7	0.01**
		KPSS	0.1012	4	0.1
	1980.09-1985.09	ADF	-3.3952	3	0.0651*
		KPSS	0.3164	1	0.1
	1985.10-1993.06	ADF	-3.2577	4	0.0832*
		KPSS	0.1796	2	0.1
	1993.07-1998.11	ADF	-3.9429	3	0.0179**
		KPSS	0.088	1	0.1
	1998.12-2003.12	ADF	-3.5106	3	0.0483**
		KPSS	0.0641	1	0.1
	2004.01-2014.12	ADF	-3.4809	5	0.0468**
		KPSS	0.1171	2	0.1

Table A2 Diagnostics tests for VAR (p) specification between each indicator and IDPs

Panel A Macro indicator constructed by all the factors								
Interval	Obs	Lag	Serial correlation		ARCH effect test		Jarque-Bera normality tests	
			Test Statistic	P-value	Test Statistic	P-value	Test Statistic	P-value
1980.09-2014.12	412	7	50.996	0.1141	163.16	< 2.67E-15	43567	< 2.20E-16
1980.09-1985.09	61	4	35.405	0.9114	38.217	0.9476	4.847	0.9176
1985.10-1993.06	93	4	53.22	0.2803	40.999	0.6421	5.7588	0.2179
1993.07-1998.11	65	4	39.263	0.8114	50.181	0.2755	2.6251	0.6224
1998.12-2003.12	61	4	46.561	0.5319	46.301	0.4184	1.1996	0.8782
2004.01-2014.12	132	4	60.097	0.113	63.297	0.03721	2166.9	< 2.2e-16

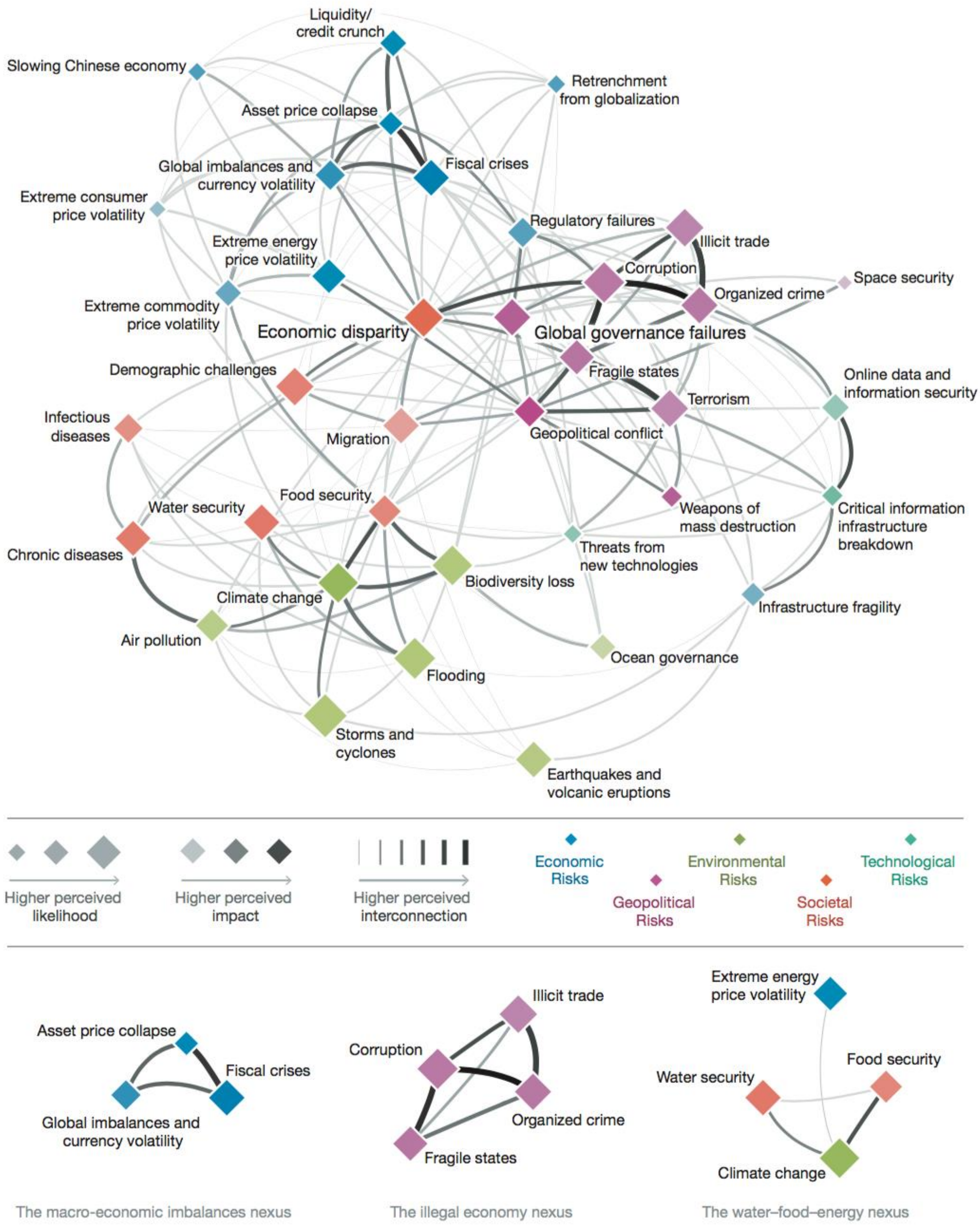
Panel B Macro indicator constructed by leading factors								
Interval	Obs	Lag	Serial correlation		ARCH effect test		Jarque-Bera normality tests	
			Test Statistic	P-value	Test Statistic	P-value	Test Statistic	P-value
1980.09-2014.12	412	4	43.111	0.673	126.13	1.28E-09	43332	< 2.2e-16
1980.09-1985.09	61	4	41.608	0.7307	44.206	0.5055	10.511	0.03265
1985.10-1993.06	93	4	38.326	0.8398	44.453	0.495	1.5773	0.8129
1993.07-1998.11	65	4	53.082	0.2847	43.479	0.5365	1.5311	0.8211
1998.12-2003.12	61	3	54.019	0.3972	60.017	0.06638	0.7674	0.9428
2004.01-2014.12	132	4	33.977	0.9371	54.019	0.3972	0.7674	0.9428

Panel C Macro indicator constructed by procyclical factors								
Interval	Obs	Lag	Serial correlation		ARCH effect test		Jarque-Bera normality tests	
			Test Statistic	P-value	Test Statistic	P-value	Test Statistic	P-value
1980.09-2014.12	412	4	61.787	0.08725	199.33	< 2.2e-16	43844	< 2.2e-16
1980.09-1985.09	61	4	46.252	0.5447	39.253	0.7131	9.726	0.04531
1985.10-1993.06	93	4	37.809	0.8543	47.018	0.3898	1.9596	0.7432
1993.07-1998.11	65	3	52.085	0.4706	44.367	0.4986	1.6882	0.7929
1998.12-2003.12	61	4	46.722	0.5252	38.438	0.7445	2.4308	0.6571
2004.01-2014.12	132	4	55.195	0.2213	58.775	0.0816	2902.5	< 2.2e-16

Panel D Macro indicator constructed by the factors in the 6 economic groups								
Interval	Obs	Lag	Serial correlation		ARCH effect test		Jarque-Bera normality tests	
			Test Statistic	P-value	Test Statistic	P-value	Test Statistic	P-value
1980.09-2014.12	412	4	62.341	0.07993	186.01	< 2.2e-16	36091	< 2.2e-16
1980.09-1985.09	61	4	38.946	0.8213	40.963	0.6436	0.71069	0.95
1985.10-1993.06	93	4	59.546	0.1226	37.061	0.794	2.6419	0.6194
1993.07-1998.11	65	4	38.455	0.836	37.063	0.7939	2.1501	0.7082
1998.12-2003.12	61	4	44.682	0.6096	36.57	0.8105	1.8179	0.7692
2004.01-2014.12	132	4	53.267	0.2788	61.909	0.04781	2065.2	< 2.2e-16

Panel E Macro indicator constructed by the effective factors extracted by Lasso								
Interval	Obs	Lag	Serial correlation		ARCH effect test		Jarque-Bera normality tests	
			Test Statistic	P-value	Test Statistic	P-value	Test Statistic	P-value
1980.09-2014.12	412	5	56.364	0.1001	115.06	4.70E-08	56152	< 2.2e-16
1980.09-1985.09	61	4	36.618	0.8848	43.329	0.5429	8.5893	0.07223
1985.10-1993.06	93	4	35.762	0.904	40.621	0.6578	1.5943	0.8098
1993.07-1998.11	65	4	44.641	0.6113	37.819	0.7674	1.8167	0.7694
1998.12-2003.12	61	4	46.088	0.5515	46.278	0.4193	2.1371	0.7106
2004.01-2014.12	132	4	61.535	0.09076	69.912	0.01009	2903.6	< 2.2e-16

Figure A1 Risks Interconnection Map 2011 illustrating systemic interdependencies in the hyper-connected world we are living in.



Source: this figure is obtained from Helbing (2013)