Journal of Asian Economics 84 (2023) 101574

Contents lists available at ScienceDirect

Journal of Asian Economics

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

Full length article

Modeling an early warning system for household debt risk in Korea: A simple deep learning approach

Yujin Kwon, Sung Y. Park *

School of Economics, Chung-Ang University, 84 Heukseok-Ro, Dongjak-Gu, Seoul, Republic of Korea

ARTICLE INFO

JEL classification:
C51
C45
E44
E58

Keywords:
Household debt
Debt ratio
Debt crisis
Early warning indicator
Deep learning

ABSTRACT

We construct an early warning indicator for household debt risk by analyzing the relationship between household debt and certain important macroeconomic determinants using a simple deep learning approach. A precise and informative indicator can help inform economic policies, especially in light of the recent growth in the ratio of household debt to income. Although several studies have analyzed the determinants of the household debt crisis, very few have examined early warning indicators for household debt risk. Some studies suggest that a situation can be regarded as a crisis if the household debt ratio is greater than 50% or 85%. However, as the household debt ratio in Korea is already over this threshold, this criterion is neither informative nor useful. Accordingly, we propose a transformed index that addresses long-term memory characteristics. Moreover, five categories for the degree of household debt crisis are considered instead of the binary variable that has been frequently used in previous studies. Furthermore, we use a well-known deep learning approach to find a non-linear relationship between crisis indices and many factors. The empirical results demonstrate that the proposed early warning indicator explains the household debt crisis quite well.

1. Introduction

The value of household debt in Korea reached approximately two quadrillion won in 2020 and continues to increase. In particular, in 2020, household debt was 106% of the nominal gross domestic product (GDP), a value much higher than the debt crisis thresholds of 50% or 85% suggested by previous studies (Arcand et al., 2015; Cecchetti et al., 2011). This steady increase in household debt is of great importance for policymakers in addressing financial vulnerability and risk at the household level.

An increase in household debt does not always signal a debt crisis. Household debt promotes economic growth and improves social welfare by smoothing economic activities (Cecchetti et al., 2011). Beyond a certain level, however, debt can have detrimental effects on economic growth due to its contractual features (Kim et al., 2014), and these effects often trigger financial crises (Allen & Gale, 2000). Therefore, it is crucial to develop an adequate early warning system (EWS) that can efficiently identify household debt risk signals and respond to the shock of further economic downturns. The EWS began to attract attention following the financial crises in the 1990s, when the EWS started to play an important role in coping with the crisis. From this perspective, EWSs can bring substantial value to policymakers by allowing them to detect underlying economic weaknesses and vulnerabilities, and possibly take preemptive steps to reduce the crisis risk.

A number of studies warn of the serious impact of household debt on the economic crisis and suggest that we should be vigilant about the economic crisis it will cause (Cynamon & Fazzari, 2008; Zabai, 2017). For this, it is essential to study the household

* Corresponding author.
E-mail address: sungpark@cau.ac.kr (S.Y. Park).

https://doi.org/10.1016/j.asieco.2022.101574
Received 18 May 2022; Received in revised form 22 November 2022; Accepted 16 December 2022
Available online 20 December 2022

debt crisis early warning model. Although an early warning of household debt is becoming increasingly important, there are very few studies on an EWS for household debt risk. Studies on EWSs for an overall financial or banking crisis account for most of the research on EWS (Qin & Luo, 2014). A classical EWS model uses a non-parametric method that can capture abnormal movements of macroeconomic variables or a parametric method using the logit or probit model (Berg & Pattillo, 1999; Kamin & Babson, 1999; Kamin et al., 2001; Kaminsky et al., 1998). Recent studies have employed different methods to improve the ability to classify crises. Duca and Peltonen (2013) propose a framework that combines a logit model with a signaling approach concerning the tradeoff between type I and type II errors. Caggiano et al. (2014) propose a multinomial logit approach that improves the predictive power of EWS compared to the binomial logit model.

The relationships between household debt and other related variables can be represented by a non-linear correspondence rather than in a linear fashion. Thus, a non-linear and parsimonious model is needed. Some recent studies have shown that forecasting models using machine learning and deep learning approaches provide reliable and superior results in terms of forecasting accuracy (Malagrino et al., 2018; Rather et al., 2015; Ticknor, 2013; Holopainen & Sarlin, 2017). Alessi and Detken (2018) as well as Tanaka et al. (2016) argue that random forests improve early warning predictions in comparison to the logit model and the signaling approach. A deep learning approach has recently been applied to some economic and financial models. Deep neural networks (DNNs) are artificial neural networks (ANNs) with multiple hidden layers. In particular, for the time-series model, there are several neural network models with recurrent neural networks (RNNs). An RNN is a special type of ANN that permits continuing information related to past knowledge by utilizing a special type of looped architecture (Cho et al., 2014; Sutskeve et al, 2014). Such RNN-based methods and their variations – including early works using naive RNN models (Connor et al., 1992), hybrid models combining the use of an autoregressive integrated moving average (ARIMA) and multilayer perceptron (Jain & Kumar, 2007; Zhang, 2003; Zhang et al., 1998), or a combination of RNN and dynamic Boltzmann machines (Dasgupta & Osogami, 2017) – have been shown to outperform non-deep learning models in time-series forecasting. Many of the advanced architectures today are inspired by RNN. The long-short term memory (LSTM) network is special type of RNN, capable of learning long-term dependencies. Siami-Namini et al. (2018) compare the predictive power of the ARIMA model with the LSTM model using several economics time series and find that the predictive power of LSTM is higher than that of the ARIMA model. Therefore, in this study, we conduct classification and prediction based on DNN and LSTM, and compare the two models to determine which model exhibits higher predictive power.

Moreover, we develop a new index with the logarithm of the debt-to-GDP ratio, which is introduced as an index in Park (2017). Considering the characters of long-term memory in the previous index, Michelacci and Zaffaroni (2000) explains long-memory process of GDP per capita, and find that it exhibit long memory, and mean-reverting. we adopt a fractionally differenced index that explains the household debt risk quite well. Song (2010) adopts a multi-level crisis indicator rather than a binary indicator, which is very limited in measuring the intensity of the crisis at several levels. Accordingly, we divide the household debt risk into five levels including the baseline, which is calculated using the rolling window method.

To construct a well-specified EWS, selecting appropriate input variables is of great importance. Some empirical studies have analyzed the determinants of household debt in many countries (Barnes & Young, 2003; Kim et al., 2017). For example, Meniago et al. (2013) find that increasing household debt can be associated to positive changes in house prices, CPI, GDP, household consumption expenditures and household savings. Stockhammer and Wildauer (2018) suggest that real estate prices are the most important driver of household debt. Moreover, recent studies show that the predictive power of a model with Internet search data is significantly higher than that of a model without such data. Galil and Soffer (2011) and Niesert et al. (2020) confirms that newspaper search keywords are closely related to the cyclical values of the leading economic and economic sentiment indexes. News reports have been analyzed in closely related literature to understand their impact on different sources of risk (Nyman et al., 2021; Smales, 2016; Tsai et al., 2016). One potential reason for using article keywords to forecast macroeconomic values is their ability to forecast future uncertainties that may arise as a result of structural changes, which point forecasts hardly capture (Rambaccussing & Kwiatkowski, 2020). In this study, we use ten macroeconomic variables that are employed as explanatory variables in Park (2017); three article search keywords that show a high correlation with household debt and the EWS risk level; and the marriage variable, which is the only microeconomic variable used in this study, based on the life cycle hypothesis.

In addition, we consider three different scenarios of the future economic situation to determine how changes in the main determinants cause household debt risks. In particular, we focus on the relationship between house prices and household debt risk. Karasulu (2008) analyzes the determinants of household debt in Korea and reports that the key factor affecting aggregate debt levels is the homeownership decision of households. Therefore, to determine how the household debt risk level reacts to the varying housing price index, three scenarios are designed in this study: (i) The house price index growth rate is 3.7%; (ii) The house price index growth rate increases from 4.2% to 7.7%; (iii) The house price index growth rate increases from 4.7% to 11.7%. The house price index growth rate in the first quarter of 2021 was 3.7%; therefore, scenario (i) assumes that the future growth rates of the house price index are the same as in the first quarter of 2021. Scenario (ii) assumes that the growth rate of the house price index rises by 0.5%p in each quarter. The last scenario (iii) assumes a sharp rise in housing prices to identify the response of household debt to soaring housing prices. This scenario assumes that the growth rate of the house price index rises by 1%p each quarter. The results of this scenario simulation suggest that an increase in the housing price index affects both short -and long-term household debt risks.

This study implements an EWS model that can predict household debt risk level relatively accurately and overcomes the limitations of existing studies. We used long memory considering indicators to differentiate this study from previous EWS studies. Furthermore, overcoming the limitations of previous EWS models that predict household debt risk using linear models, such as traditional signaling approaches and logit or probit models, we propose a non-linear model with more accurate time-series predictions by applying a deep learning approach.

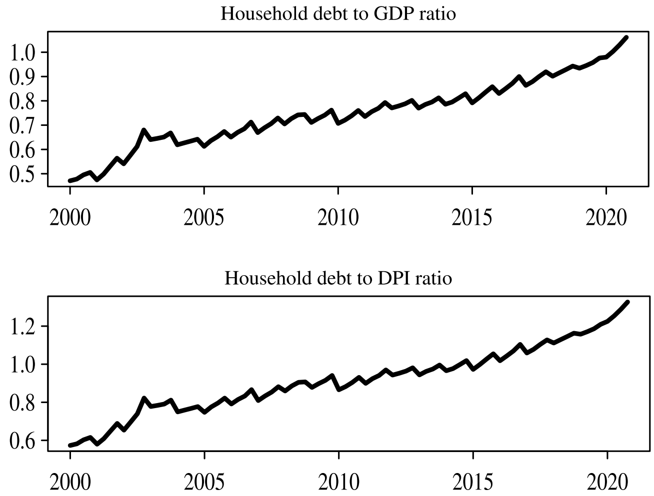

Fig. 1. Debt-to-GDP and Debt-to-DPI ratio.
Source: National Accounts, Bank of Korea: www.ecos.bok.or.kr.

The remainder of the paper is organized as follows: the models used to estimate the index and predict the future index are described in Section 2. Section 3 explains the data and descriptive statistics and summarizes the empirical results. This is followed by a conclusion and discussion in Section 4.

2. Model

2.1. Indicator of household debt

Household debt in Korea consists of household loans and sales credit from depositors and other financial institutions. Korea's household debt began to increase in the early 2000s and has accumulated continuously over the past 20 years, raising concerns about economic resilience as the household debt-to-GDP ratio is relatively large compared to some other countries. Korea's household debt, which has been increasing since 2013, peaked in the second half of 2016 and then declined, but it has continuously increasing trend. Although the growth rate of household debt has slowed, the household debt repayment burden has increased as the growth rate of household debt still outpaced the growth rate of income and financial assets.

The ratio of household debt to GDP as well as to gross disposable income (DPI) has shown an increasing trend since the 2000s (see Fig. 1). Particularly, the debt-to-GDP ratio was larger than 100% in 2020.

However, it is challenging to determine whether a situation is dangerous based simply on the debt-to-GDP or debt-to-DPI ratios. In previous studies, the debt-to-GDP ratio was used as an index to detect a household debt crisis. Cecchetti et al. (2011) estimate that the threshold at which the debt-to-GDP ratio hampers economic growth is 84%, but this threshold is not statistically significant. Arcand et al. (2015) estimate a threshold of 50%, using data of 120 countries from 1960 to 2010. However, the domestic debt-to-GDP ratio in Korea exceeded 50% in 2001 and surpassed 106% in 2020. Therefore, it is difficult to clearly define a threshold that represents an actual crisis in light of this increasing trend of household debt.

Moreover, an increase in household debt does not always have a negative effect on economic growth. It can also have a positive effect on total consumption, which in turn promotes economic growth. Therefore, it may not be desirable to use the raw time series of the debt-to-GDP ratio to identify the timing and level of the crisis. Rather, a suitable indicator of household debt risk is a time series of the growth rate of GDP subtracted from that of household debt, which can be derived by comparing the growth rates of GDP and household debt. A simple index can be defined as in (2.1).[1]

$$Index_1 = \Delta \log(HD_t) - \Delta \log(GDP_t). \tag{2.1}$$

However, as Fig. 2 demonstrates, $Index_1$ may not adequately capture the household debt risk. While the debt-to-GDP ratio is steadily increasing, $Index_1$ does not show an increasing trend. Kang (2017) suggests that household debt has both flow and stock effects. Flow effects promote economic growth, while stock effects have the opposite effect. Therefore, these two characteristics need to be considered in the index of household debt risk. However, $index_1$ only explains the flow effect of household debt. To deal with this problem, Park (2017) uses a convex combination of the flow and stock data of $\log(HD_t/GDP_t)$. However, the method

[1] $Index_1$ can be also directly calculated by the first differences of $\log(HD_t/GDP_t)$.

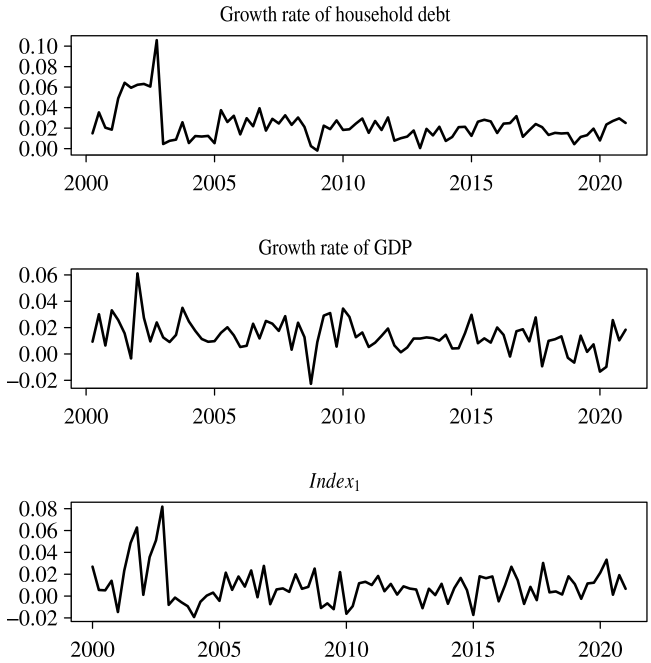

Fig. 2. Growth rate of household debt and $Index_1$.

of selecting the weight in that study is quite ad hoc. To consider both the long- and short-term effects of $\log(HD_t/GDP_t)$, a more plausible index is required.

To construct an appropriate index for household debt risk that has an increasing trend with distinct fluctuations, we use the fractional difference of $\log(HD_t/GDP_t)$ instead of the first difference. If the time series data are not an integrated time series with order 1, differentiating with order 1 may cause a loss of information. To minimize this loss, an optimal order for the difference is required.

Granger (1980) and Hosking (1981) start to use fractional difference, and explain it with the definition of "long memory". There are several definitions of long memory; in this study, we used long memory with a discrete time series process y_t and an autocorrelation function ρ_j at lag j. According to McLeod and Hipel (1978), a process possesses long memory if the quantity

$$\lim_{n\to\infty} \sum_{j=-n}^{n} |\rho_j| \tag{2.2}$$

is non-finite. Granger and Joyeux (1980) and Hosking (1981) also propose the autoregressive fractional integrated moving average model (ARFIMA). Assuming $\{y_t\}, t = 1, 2, \dots, T$, is an ARFIMA (0, d,0) model that has the following equation of discrete time stochastic process

$$(1 - L)^d y_t = \epsilon_t, \tag{2.3}$$

where L is the lag operator and d is the order of the fractional difference. Diaz and Osler (1974) define the order of the fractional difference as any real or complex number. ϵ_t is the white noise process of the random variables. Samko and Ross (1993) show that the expression $(1 - L)^d$ can be defined by the series expansion in the following form

$$(1 - L)^d = \sum_{m=0}^{\infty} (-1)^m \binom{d}{m} L^m, \tag{2.4}$$

where $\binom{d}{m}$ are the binomial coefficients that are defined by the following equation[2]

$$\binom{d}{m} = \frac{\Gamma(d+1)}{\Gamma(d-m+1)\Gamma(m+1)} = \frac{(-1)^{m-1} d\,\Gamma(m-d)}{\Gamma(1-d)\Gamma(m+1)}. \tag{2.5}$$

and (2.4) can be written as

$$\Delta^d = (1 - L)^d \tag{2.6}$$

[2] See the more detailed Eq. (2.5) in Samko and Ross (1993).

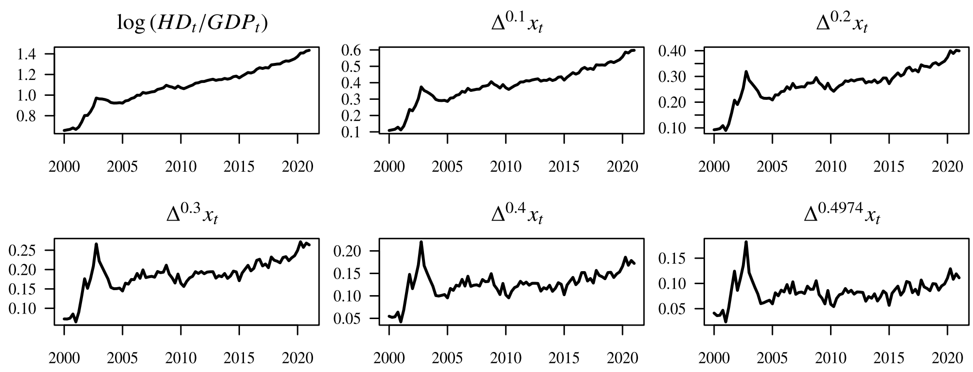

Fig. 3. Fractionally differenced time series.
Notes: $\Delta^n x_t$ denotes the nth order difference, and 0.4974 in the last graph is the maximum likelihood estimate. For simplicity, x_t denotes $\log(HD_t/GDP_t)$.

$$= 1 - dL - \frac{1}{2}d(1-d)L^2 - \frac{1}{6}d(1-d)(2-d)L^3 - \cdots$$
$$= \sum_{m=0}^{\infty} \frac{\Gamma(m-d)}{\Gamma(-d)\Gamma(m+1)} L^m.$$

This operator is called the Grunwald–Letnikov fractional difference of order d with the step T=1.[3] One can easily check that this model becomes an AR(1) model when d equals 1.

To consider the long-memory characteristics of $\log(HD_t/GDP_t)$, we defined $index_2$ as

$$Index_2 = \Delta^d \log(\frac{HD_t}{GDP_t}). \tag{2.7}$$

The fractional differencing process can be regarded as a combination of the $I(0)$ and $I(1)$ paradigms. Caggiano and Leonida (2009) tests for convergence in international output, and suggest that forcing the data to be either $I(1)$ or $I(0)$ is likely to be misleading when examining convergence properties of different economies, $I(d)$ processes can be correct understanding of transitional dynamics. Fig. 3 shows the fractionally differenced time series of $\log(HD_t/GDP_t)$. The fractionally differenced time series have more mean-reverting tendency than the raw time series. Furthermore, the smaller the value of d, the more prominent the trend is in the fractionally differenced time series, and the longer the memory of the values of previous periods. However, as d increases, the trend is less evident in the fractional time series, and the long memory of the values of previous periods is weakened. Therefore, the selection of parameter d is crucial.

Bollerslev and Mikkelsen (1996) show that the series displays a long-term memory if $0 < d < 0.5$ with a mean-reverting property. If d is $0.5 < d < 1$, the series still displays long-term memory, but the weights it assigns to past values are very small. Therefore, we consider the case over the interval $(0, 0.5)$. The optimal fractional differencing parameter can be estimated by applying the maximum likelihood estimation method (Haslett & Raftery, 1989).[4] The maximum likelihood estimate is 0.4974.[5]

Using $Index_{2,t}$, we construct an EWS index to identify household debt risk. For this, we use the rolling median and rolling average methods, and we define them as follows:

$$\gamma_t = Median(\{Index_{2,t-20}, \ldots, Index_{2,t-1}\}), \tag{2.8}$$

$$\delta_t = \frac{Index_{2,t-20} + \cdots + Index_{2,t-1}}{20}, \tag{2.9}$$

where t denotes quarters. Both methods use the preceding 20 quarters as the rolling window. By employing a time-varying index using a rolling window, we avoid abrupt changes in the EWS index. We set the rolling median or mean as the baseline as the point at which there is no economic impact.

The next question is whether to use a crisis indicator that varies with the severity of the crisis or whether to adopt the 0/1 dichotomy. However, binary indicators are very limited in measuring the intensity of a crisis at several levels. Thus, in this study, we categorize the EWS index into five levels: $x = 1; 2; 3; 4; 5$ and $y = 1; 2; 3; 4; 5$, where x is the risk level calculated with the rolling median baseline and y is the risk level calculated with the rolling average baseline. The five different crisis levels with the rolling median baseline are determined according to the following specification:

$$x = 1 \quad if \qquad Index_{2,t} < \gamma_t, \tag{2.10}$$

[3] The Grunwald–Letnikov fractional difference Δ^{α}_{T} of order α with the step T is defined by the equation $\Delta^{\alpha}_{T} y(t) = (1 - L_T)^{\alpha} y(t) = \sum_{m=0}^{\infty} (-1)^m \binom{d}{m} y(t - mT)$ (Scherer et al., 2011).

[4] See the detailed procedure for maximum likelihood estimation in Haslett and Raftery (1989).

[5] The standard error of the parameter estimates is 0.0000018.

Table 1
Thresholds and bounds.

Risk level	Number of observations	
	Rolling median threshold	Rolling average threshold
1	32	37
2	13	30
3	16	8
4	14	8
5	10	2

Notes: Level 5 in the rolling average threshold includes 2002 Q4 and 2020 Q2.

$$
\begin{aligned}
x = 2 \quad if \quad & \gamma_t \leq Index_{2,t} < \gamma_t^{q_1}, \\
x = 3 \quad if \quad & \gamma_t^{q_1} \leq Index_{2,t} < \gamma_t^{q_2}, \\
x = 4 \quad if \quad & \gamma_t^{q_2} \leq Index_{2,t} < \gamma_t^{q_3}, \\
x = 5 \quad if \quad & \gamma_t^{q_3} \leq Index_{2,t},
\end{aligned}
$$

where γ_t^q indicates the q-quantile of $\{Index_{2,t-20}, \ldots, Index_{2,t-1}\}$. The different risk levels with the rolling average baseline are determined according to the following specification:

$$
\begin{aligned}
y = 1 \quad if \quad & Index_{2,t} < \delta_t, \\
y = 2 \quad if \quad & \delta_t \leq Index_{2,t} < \delta_t + c_1 \cdot s.d, \\
y = 3 \quad if \quad & \delta_t + c_1 \cdot s.d \leq Index_{2,t} < \delta_t + c_2 \cdot s.d, \\
y = 4 \quad if \quad & \delta_t + c_2 \cdot s.d \leq Index_{2,t} < \delta_t + c_3 \cdot s.d, \\
y = 5 \quad if \quad & \delta_t + c_3 \cdot s.d \leq Index_{2,t},
\end{aligned}
\tag{2.11}
$$

where y denotes the EWS risk level, and $s.d$ is the rolling window standard error of the rolling average. We note that if the index movements become volatile, $s.d$ increases. This causes the bandwidths of the risk levels to rise, which implies that abrupt movements of the index do not change the levels quickly. Therefore, the use of $s.d$ is a plausible method to adequately divide risk levels. However, for the rolling median method, because each rolling standard deviation is too small to divide the risk levels proportionally, $s.d$ cannot be used as an appropriate method. In this case, risk levels with quantiles can be considered. Since each quantile is a cut point dividing the range of a probability distribution, it is a good method to classify each the risk level.

In (2.10) and (2.11), q_i or c_i for $i = 1, 2, 3$ decides the bandwidths of adjacent risk levels. Here, we encounter the issue of arbitrariness, as there are no rules for measuring the severity of household debt risk. In this study, we set $q_1 = 0.6$, $q_2 = 0.8$, and $q_3 = 0.99$ for the quantile level, this quantile level is selected to keep the EWS conservative, so that the EWS does not issue a risk warning of level 4–5 too often. Also we set $c_1 = 1, c_2 = 1.64$, and $c_3 = 2.33$, as these are the z-scores at the 85%, 90%, and 98% confidence intervals, respectively.

To select the adequate baseline and bounds, we define well-divided risk levels based on two criteria: (i) Level 5 includes the 2002 credit card lending distress; (ii) The indicator can capture the current household debt risk situation. As we can see in Table 1 and Fig. 4, the rolling median threshold and its bounds do not match the criteria. Even though there are many observations at risk level 5 in the current situation, it cannot capture the 2002 credit card lending distress at risk level 5. Since the rolling median is robust to anomalies, the sudden risk can be ignored by smoothing. By contrast, the rolling average threshold matches the criteria. The index warns of a high crisis level in recent times and adequately captures the 2002 credit card lending distress at risk level 5. Accordingly, we used $Index_2$ with a rolling average baseline, and each bound was calculated using its rolling standard deviation.

2.2. Deep learning model

A DNN is an ANN with multiple hidden layers (usually, at least more than two hidden layers). In addition to multiple layers, it has numerous activation functions and gradient descent optimizers that help overcome the problem of ANNs. By adding hidden layers and non-linear activation functions, the model can capture more complex relationships in the data than the ANN can. One of the major problems with ANNs is the vanishing gradient problem, which arises when the network is trying to learn a model, but the gradients of an error are so small that adjustments to the weights through backpropagation make no difference to the learning process and a global minimum is never reached (Maxwell et al., 2017). However, specifically for deep learning, instead of typical general activation functions – such as sigmoid or hyperbolic tangent functions – a different set of activation functions is used to achieve better results. One of these activation functions is a rectified linear unit (ReLU). The ReLU function is expressed as follows:

$$f(x) = max\,(0, x), \tag{2.12}$$

where x is the result of the equation coming from the node of the network. The ReLU function does not deal with an evaluation that is below zero, and the value can only be between 0 and 1. We use the ReLU activation function in DNN to solve the vanishing gradient problem.

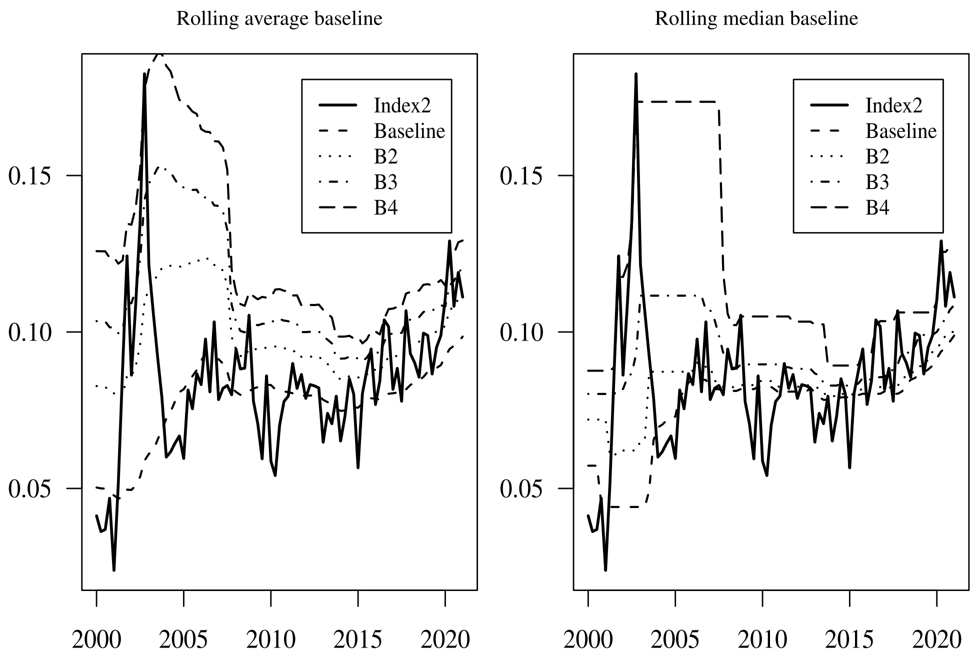

Fig. 4. Thresholds and bounds.
Notes: Each rolling window uses values of the past 5 years to set the baseline. The level below the baseline is risk level 1. The level between the baseline and B2 is risk level 2, followed by levels 3 and 4. The B5 line indicates risk level 5.

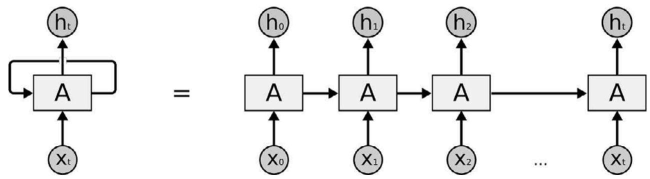

Fig. 5. Basic structure of the RNN. Notes: x_t is the input value, and h_t is the hidden layer of time t.
Source: colah.github.io

The output layer relies on the cost and softmax functions. The softmax function can be defined as follows:

$$\text{softmax}(\mathrm{n})_i = \frac{e^{n_i}}{\sum_{k=1}^{K} e^{n_k}} \text{ for } \ i = 1 \ldots K, \tag{2.13}$$

where n is the result of the equation coming from the node of the last layer, and K is the number of nodes. The idea of the softmax function is straightforward. By normalizing the data such that the output layer value lies in the range (0, 1), the sum of the output layer values is equal to 1. These can then be interpreted as probabilities, where the highest probability is most likely the best candidate label for classification.

However, it is also well known that the deep learning model has an overfitting problem. Overfitting occurs when the model provides a near-perfect fit for the in-sample but poor predictions for the out-of-sample. To deal with overfitting problem, setting an optimal hyperparameter is crucial when constructing a deep learning model. Through the restricted Boltzmann machine, a weight initialization method devised by Hinton et al. (2006), active research is being conducted on DNNs.[6] Furthermore, by randomly excluding the nodes during training, the model can offer effective regularization to reduce overfitting. This process is called a dropout, and we can set the hyperparameter α, which is the proportion of dropping nodes in the learning model. Moreover, early stopping can be used in learning to avoid over-training and hence overfitting the model. Accordingly, we adopt early stopping in this study. An over-trained model has a tendency to memorize all the training data points. However, with early stopping, a large number of training epochs is specified. The model is stopped from training further when its performance with the validation dataset stops improving. In this study, we use dropout and early stopping for all trained models to obtain optimal results.

[6] See more details about weight initializing in Hinton et al. (2006).

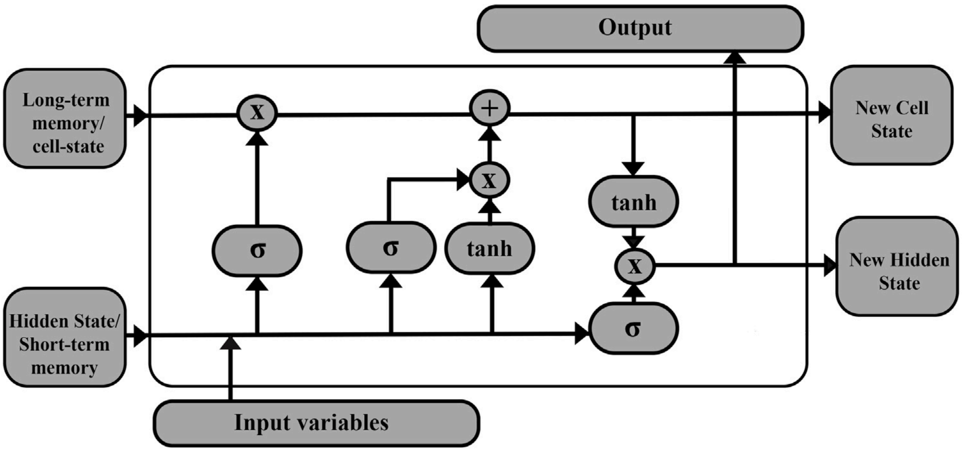

Fig. 6. Basic structure of LSTM. Notes: Input variables are macroeconomic variables. X is multiplication sign, and + indicates plus sign.

Unlike general ANNs, RNNs can effectively model sequential or time-series information because the hidden layer vector h, which is updated at time step t, is considered when outputting the next input value (Pascanu et al., 2013). The basic structure of the RNN is shown in Fig. 5, and the hidden layer vector h can be calculated with this activation function as follows.

$$\begin{aligned} h_t &= \tanh(Wh_{t-1} + Ix_t) \\ &= \frac{e^{Wh_{t-1}+Ix_t} - e^{-(Wh_{t-1}+Ix_t)}}{e^{Wh_{t-1}+Ix_t} + e^{-(Wh_{t-1}+Ix_t)}}. \end{aligned} \tag{2.14}$$

In RNNs, unlike DNN, we use a different activation function called $\tanh(\cdot)$. The activation function $\tanh(\cdot)$ is a hyperbolic tangent activation function, which is a rescaling of the logistic sigmoid, such that its output ranges from −1 to 1. The output of the ReLU function used in DNN ranges from 0 to 1, whereas the output of the hyperbolic tangent function used in RNN is different in that it ranges from −1 to 1. W denotes the recurrent weight matrix, I is a projection matrix, and x_t is the input value. The hidden state h_{t-1} is used to make a prediction at time t. The output of the RNN network at time t is represented as

$$y_t = \text{softmax}(Wh_{t-1}), \tag{2.15}$$

where softmax(·) is the function for classification, shown in (2.13). Using h^l as the input to another RNN, we can stack RNNs, creating deeper architectures to develop a deep learning model as follows.

$$h_t^l = \sigma(Wh_{t-1}^l + Ih_t^{l-1}), \tag{2.16}$$

where $\sigma(\cdot)$ is a logistic sigmoid function, W is a weight matrix, and l is the layer. The activations from each time step t are stored in the internal state of the network to provide a temporal memory property.

However, a major weakness of the RNN is the long-term memory requirement. As past information is continuously accumulated, old information disappears, and recent information is reflected. To overcome this drawback, Hochreiter and Schmidhuber (1997) developed the LSTM algorithm and suggested that the LSTM addresses the vanishing gradient problem commonly found in ordinary RNNs by incorporating gating functions into their state dynamics. Here, the vanishing gradient problem is the same concept as previously described in ANN, and it means a problem in which the gradient of the loss function approaches 0 as more layers are added to the neural network. When the slope of the loss function approaches zero, it becomes difficult to properly train the network because the weights cannot be adjusted through backpropagation. The basic architecture of the LSTM is shown in Fig. 6. Each LSTM block contains a memory cell, an input gate, a forget gate, and an output gate, and at each time step t, an LSTM maintains a hidden state and a cell state that controls the extent to which a value remains in its memory. Graves (2012) show the computation of LSTM, which is defined as follows:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \tag{2.17}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \tag{2.18}$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c), \tag{2.19}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t, \tag{2.20}$$

$$o_t = \sigma(W_0[h_{t-1}, x_t] + b_0), \tag{2.21}$$

$$h_t = o_t \times \tanh(C_t), \tag{2.22}$$

where $\sigma(\cdot)$ and $\tanh(\cdot)$ are the non-linear activation functions, x_t is the input value, h_t denotes the hidden value at time step t, o_t is the output value at time step t, b is a bias, W denotes the weights, and f_t is the forget gate that decides which information to discard.

Table 2
Variable descriptions.

	Variables	
Household debt	hdebt	Household and non-profit organization
GDP (season adjustment, nominal)	ngdp	Gross domestic product, nominal
GDP (season adjustment, real)	rgdp	Gross domestic product real
KOSPI index (1980.1.4=100)	kospi	KOSPI
Consumer price index (2015=100)	cpi	Nationwide total index
Household loan interest rates (%)	hir_mean	Total average
	hir_house	Household debt
Saving (%)	recep	bank and non-banking organization
Call rate (%)	call_unse	Unsecured call rate
Capital adequacy ratio (%)	bis_asset	BIS capital adequacy ratio
Sub-standard ratio (%)	loan_below	Sub-standard ratio
House price index (2019.01=100)	hidx	House price index
Unemployment (%)	unemp	Unemployment rate
Total marriage	marriage	Nationwide total marriage
News search2	news2	Search word "household debt"
News search3	news3	Search word "credit loan"
News search5	news5	Search word "credit loan interest rate"

Notes: All the macroeconomic variable data are obtained from Bank of Korea and Korea National Statistical Office, and News search variable data are obtained from BIG KINDS. The unit of household debt and GDP is billion won; the unit of household loan interest rates, savings, call rate, capital adequacy ratio, sub-standard ratio, and unemployment is %.

Similarly, i_t and $\tilde{C}_t$ are input gates with activation functions $\sigma(\cdot)$ and $\tanh(\cdot)$ that decide what information to store in the cell state, which is responsible for long-term memory. Through simple computation, the cell state is updated with the current information. Next, by running a sigmoid layer, o_t determines the parts of the cell state to output. Finally, the cell state is put through $\tanh(\cdot)$ and multiplied by the output of the sigmoid gate. In this study, x_t is macroeconomic variables at time step t, and h_{t-1} is the short-term information at time step $t-1$. So through (2.17), we decide which past information to discard. Next through (2.18) and (2.19), we decide how much new information of input variables at time step t will be used. So from past cell state C_{t-1}, new cell state is updated with linear function of discarded past information and part of new information about the macroeconomic variables. Finally, with the new information of cell state, the model output the information about which EWS index to be updated through (2.21) and (2.22); thus, by mixing long- and short-term memory operations for several time steps, the hidden layer vector h is updated. We train the most basic DNN and LSTM, and find the optimal model for an EWS for household debt crisis.

3. Empirical analysis

3.1. Data

To forecast the future EWS index, 14 factors are chosen, as summarized in Table 2. The data, except for household debt and GDP, are obtained for a period of 21 years from the 1st quarter of 2000 to the first quarter of 2021. Household debt and GDP data are collected quarterly over 26 years, from 1996 to 2021. In general, many studies analyze the determinants of household debt using micro data by household, such as labor panel data, household financial welfare survey panel data, or CB data. However, since these micro data are year-by-year data, there are limitations in estimating and predicting household debt risk. Park (2017) divided the macroeconomic variable data used in the analysis of the determinants of household debt into three types: factors affecting demand, factors affecting supply, and factors affecting both supply and demand. Among the factor variables affecting demand, real GDP growth rate, change in household loan interest rates, and growth rate of consumer price were considered. As for factor variables affecting supply, the growth rate of financial institution receipts, sub-standard loan ratio, and BIS capital adequacy ratio were considered. The stock price index and the housing price index were used as factor variables influencing all of them. Total marriage is also considered in here as an important factor based on the life-cycle hypothesis.

Recent studies show that the predictive power of a model with Internet search data is significantly high. Rambaccussing and Kwiatkowski (2020) confirms that online search behavior is a relatively reliable gauge of an individual's behavior, and that newspaper search keywords are closely related to the cyclical values of the leading economic and economic sentiment indices. In this study, we included newspaper search keywords in the set of determinants. The advantage of using article keywords for forecasting macroeconomic values is their ability to forecast future uncertainties that may arise as a result of structural changes, which point forecasts hardly capture. The data were sourced from the Bank of Korea and Korea National Statistical Office (KOSAT); and the news search word data were sourced from BIG KINDS, which provides statistics regarding Korean Internet news search keywords.[7] For the news search keywords, three main keywords are selected among 10 keywords related to household debt based on the correlation between the keywords and household debt, Index, and EWS index. Table 3 shows the correlations of each variable with household debt, Index2 defined in Eq. (2.7), and five EWS risk level. As one can see in Table 3, the keywords that show a high correlation between household debt and the EWS index are "household debt", "credit loan", and "credit loan interest rate".

[7] Bank of Korea: https://ecos.bok.or.kr, KOSAT: https://kostat.go.kr, BIG KINDS: https://www.bigkinds.or.kr.

Table 3
Correlations between selected news keywords with variables and indices.

Variables	Debt	Index	EWS index
Search word "household debt"	0.209	0.569	0.489
Search word "credit loan"	0.607	0.113	0.226
Search word "credit loan interest rate"	0.541	0.088	0.201
KOSPI index	0.877	0.102	−0.010
Consumer price index	0.961	0.179	0.118
Household loan interest rates (average)	−0.914	−0.292	−0.194
Household loan interest rates	−0.887	−0.300	−0.166
Savings	0.997	0.255	0.234
Call rate	−0.872	−0.180	−0.165
Capital adequacy ratio	0.895	0.132	0.026
Sub-standard ratio	−0.539	−0.483	−0.153
House price index	0.910	0.268	0.119
Unemployment	−0.011	−0.403	−0.130
Total marriage	−0.472	−0.154	−0.153

Notes: Variables' names are the same as in Table 2. Debt denotes Household debt, Index is Index 2 defined in Eq. (2.7), and EWS index is five EWS risk level.

Table 4
Deep learning model.

Model	Number of hidden layers	Number of lags
Deep Feed-forward Neural Network	2	2
(DNN)	3	2
Long short term memory networks	2	4
(LSTM)		6
		8
	3	4
		6
		8

Notes: The unit of the lags is a quarter. For DNN, maximum lags are denoted, since the model includes different lags which shows more accurate results than models without lags by variables.

Abbas (2005, 2017) confirms that a neural network model including several lags shows more accurate results than models without lags. However, LSTM uses its memory cell while learning; thus, there is no need for arbitrary add lags of variables. Thus, we added lags for variables only for a DNN without any memory cell in the model, and a total of 26 input variables including the lag variables were considered.[8]

3.2. Model training and forecasting

Table 4 summarizes the considered models and the number of parameters. In the deep learning model, it is important to determine the optimal number of hyperparameters. In particular, in the LSTM model, choosing an adequate lag is crucial because its choice significantly affects the results. Therefore, in consideration of the business cycle of macroeconomic variables, the models were trained with various lags, and among them, the three lags with the best predictive power were used. In this study, starting from a small number of hidden layers and lags, a total of eight models are trained. Furthermore, there is no difference in the results with learning rates of 0.01 and 0.001; therefore, a learning rate of 0.01 is used for the training. In addition, there is no difference between input nodes 52 and 104 in the DNN, and between nodes 14 and 28 in the LSTM. Thus, we use 52 nodes for DNN and 14 nodes for LSTM for simplicity. The number of trainings, 150, is the same for all models.

To address the issue of overfitting, we adopt early stopping. The early stopping method divides the data into a training set and a validation set, and uses the cross-validation method to test the validation data at each epoch. Both training loss and validation loss decrease until overfitting occurs. However, when overfitting occurs, the training loss decreases while the validation loss increases. So, early stopping forces the model to stop learning when the validation loss increases. In this study, the total number of observations is very small, and the data are imbalanced. Over the total observations of the risk, there are only two at risk level 5. Therefore, dividing the data into a training set and a test set may cause biased results.[9] Therefore, early stopping, which automatically divides the data into training and validation sets, is used for this study.

[8] Adequate variables and lags are finally selected with simple regression by ordinary least squares estimation. The selected variables are statistically significant at a p value of less than 0.1, and are as follows: unemp, cons, marriage, news2, news3, news5, unemp(L1), cons(L1), marriage(L1), news2(L1), news3(L1), cons(L2), marriage(L2), log(cpi), hir_mean, hir_house, loan_below, call_unse, log(hidx), hir_mean(L1), hir_house(L1), loan_below(L1), log(hidx)(L1), log(kospi)(L1), call_unse(L2), and loan_below(L2), where L1 and L2 denote the preceding one and two quarters, respectively.

[9] The method of 80% training set, 20% test set was also adopted, but it never reflected a risk level of 4 or 5.

Table 5
Model selection.

Models	Hidden layers	Lags	Loss	Accuracy
DNN	2	2	1.955	50.00%
	3	2	1.786	37.50%
LSTM	2	4	0.386	85.38%
		6	0.269	88.36%
		8	0.125	95.65%
	3	4	0.874	61.03%
		6	1.003	57.76%
		8	0.305	89.67%

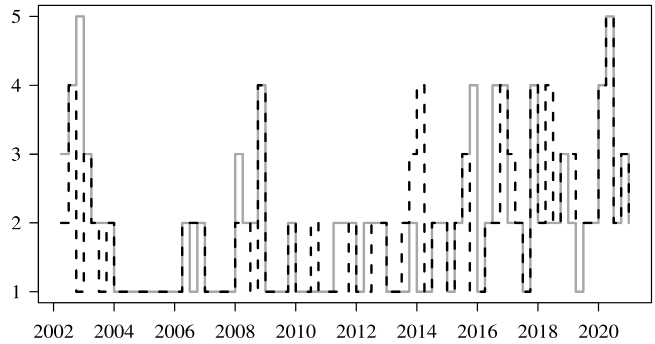

Fig. 7. Prediction of EWS index with trained model.

All factors are normalized using $(x-\min(x))/(\max(x)-\min(x))$. Normalization assigns equal weights or importance to each variable so that no single variable steers model performance in one direction just because it has a bigger weight. Furthermore, to conduct multi-class classification, all trained EWS indices are encoded into 0 or 1 via one-hot encoding. 'The one-hot encoding' is a vector encoding that converts the size of a set into the dimension of a vector, assigns a value of 1 to the element to be expressed, and assigns a value of 0 to other elements. In this paper, there are five EWS indices, and this EWS index set becomes a vector with five columns through one-hot encoding. If the EWS index is 1, a vector is created where only the first column is 1 and the other columns are 0. By doing this, we can reduce the amount of memory needed to process the data and make the model train faster and more accurate.

The loss is calculated on training and validation, and indicates how well the model is performing. In this model, we use the categorical cross-entropy loss function, which is used in multi-class classification tasks and is defined as follows:

$$CE = -\sum_{i=1}^{K} t_i \log\left(\sigma(n)_i\right), \tag{3.1}$$

where $\sigma(n)_i$ is softmax(·), as in (2.13). Thus, the categorical cross-entropy loss is multiplied by t_i with softmax(·), where t_i is only one element of the target vector, which is not zero. The target vector indicates the vector obtained through one-hot encoding. For example, if the EWS index is 1, $t_1 = 1$, and t_2, t_3, t_4, t_5 are all 0. By multiplying the log value of the softmax output by each elements in target vector, the categorical cross-entropy loss can be obtained. As a result, after several iterations, the train model reduces the loss and the model becomes more accurate. An accuracy metric is used to measure the performance of the algorithm in an interpretable manner.[10] The accuracy of a model is usually determined after the model parameters and is calculated in the form of a percentage.

Two deep learning models, DNN and LSTM, are implemented and we measure the loss and accuracies in each model. As shown in Table 5, the LSTM model can explain the training model better than DNN can. Overall, the loss in LSTM is smaller than the loss in DNN, and the accuracy of LSTM is higher than that of the DNN. Moreover, models with two hidden layers are more accurate than those with three hidden layers. In particular, the LSTM model with two hidden layers and eight lags had the least loss and the highest accuracy. Therefore, based on the results, the LSTM model with two layers and eight lags is selected to forecast the future EWS index.

We trained the model from the second quarter of 2000 to the first quarter of 2021. Fig. 7 shows the results of the trained LSTM model with eight lags. The model captures one level 5 risk in 2020, and it provides a warning of the 2002 credit card lending distress, 2007 financial crisis, and recent household debt risk quite well.

Fig. 8 shows the forecasts of the future EWS index from the second quarter of 2021 to the first quarter of 2023. The out-of-sample forecast in Fig. 8 is performed using the same variables used when training the model. For out-of-sample forecast, 10 macroeconomic

[10] See the more detail about cross entropy loss for training deep neural networks in Zhang and Sabuncu (2018).

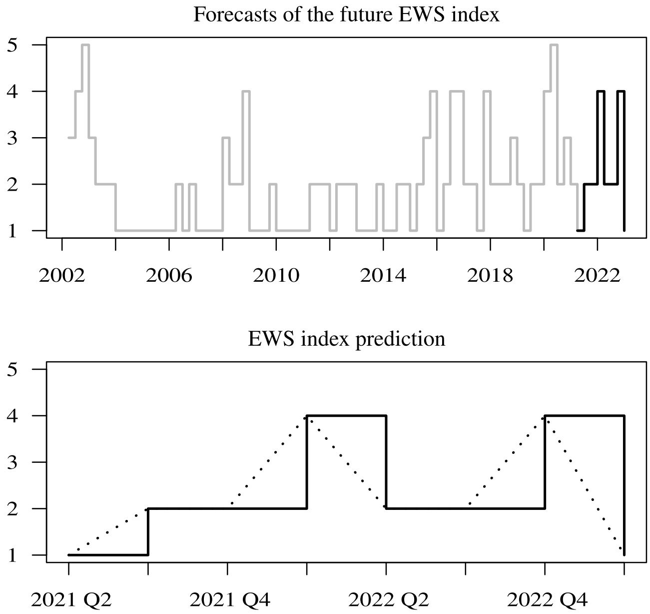

Fig. 8. EWS index forecasts. Notes: The forecast index is from the second quarter of 2021 to the first quarter of 2023.

variables presented in Table 3, 3 newspaper keyword variables, and marriage variables are used. Using the model trained within the sample, we predicted the EWS non-sampling index from second quarter of 2021 to first quarter of 2023. Therefore, the model predicted the EWS index from the second quarter of 2021 to the first quarter of 2023. The trained model predicts risk level 1–2 from the second quarter of 2021 to the fourth quarter of 2021. However, from the first quarter of 2022, the risk level rises to level 4, and the last quarter, 1st quarter of 2023, the forecast value is still returning to risk level 4. These frequent instances of high risk suggest that we should continue to be concerned about household debt and seek ways to lower the risk level of household debt.

3.3. Scenario simulation : House price changes

We focus on the relationship between house price and household debt risk. Stockhammer and Wildauer (2018) suggest that real estate prices are the most important drivers of household debt. Kim et al. (2014) find that asset price hikes contribute to the increasing trend of household debt. Karasulu (2008) identifies households' homeownership decision as the main determinant of soaring household debt in Korea. Furthermore, recently, rising housing prices coupled with low interest rates have increased the demand for mortgages, leading to an increase in household debt. Therefore, to determine how the household debt risk level reacts to the varying housing price index, the three scenarios were considered. The scenarios deal with eight quarters of prediction, and it starts from the second quarter of 2023, ends at the first quarter of 2025. The three scenarios were designed as follows.

(i) The house price index growth rate is fixed at 3.7% for eight quarters.

(ii) The house price index growth rate increases from 4.2% to 7.7% for eight quarters.

(iii) The house price index growth rate increases from 4.7% to 11.7% for eight quarters..

The house price index growth rate in the first quarter of 2021 was 3.7%, and each scenario respectively assumes a 0%p, 0.5%p, 1%p increase in the growth rate each quarter; therefore each scenario starts with 3.7%, 4.2%, 4.7% respectively.[11] We assume that the other variables, except for the house price index, following the conditional forecasting results of the ARIMAX model. Fig. 9 shows the forecast values of the macroeconomic factors with three news search variables.

First figure of Fig. 10 shows the predicted EWS index with the variable forecasts with no scenario. From 2023 to 2024, the risk level is almost fixed at risk level 2, and two more blocks from the last quarter. For the first scenario, we changed the house price index growth rate to 3.7%.

However, the results of the predicted EWS index in scenario 1 are the same as the results with no scenario which can be seen in the first figure in Fig. 10, and from 2023 to 2024, the risk level is almost fixed at risk level 2. From the first quarter of 2024 to the second quarter of 2024, the EWS index reaches risk level 3, but returns to risk level 2 subsequently.

For the second scenario, we changed the house price index growth rate from 4.2% to 7.7%, which assumes that the growth rate of the house price index rises by 0.5%p each quarter. Second figure of Fig. 10 shows the results of Scenario 2. The risk level reaches

[11] The house price index growth rate in the first quarter of 2021 was 3.7%.

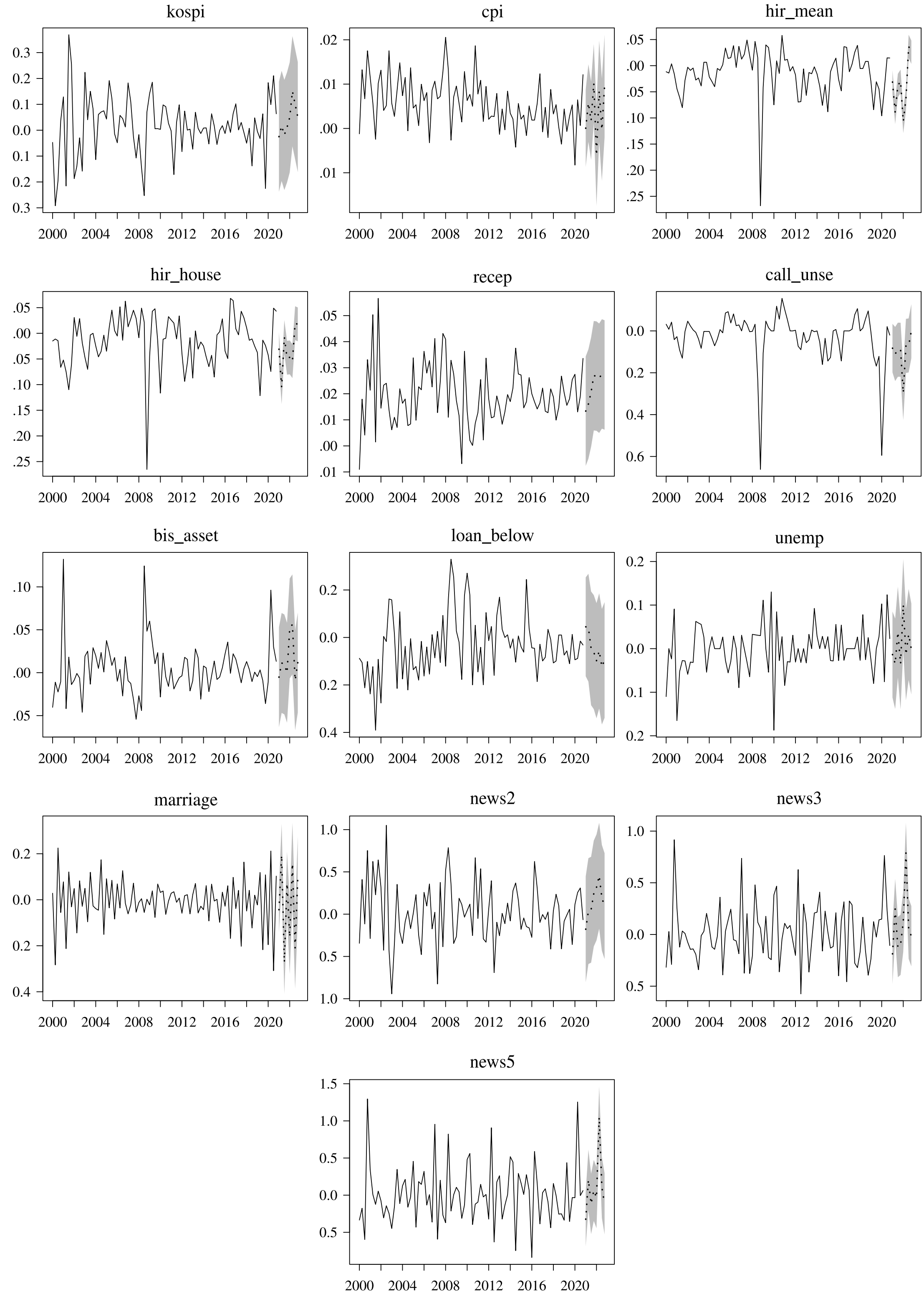

Fig. 9. Forecasts of factors with ARIMAX model.
Notes: Eight-quarter-ahead forecast values are denoted by dotted lines. The shaded area denotes the 95% confidence interval.

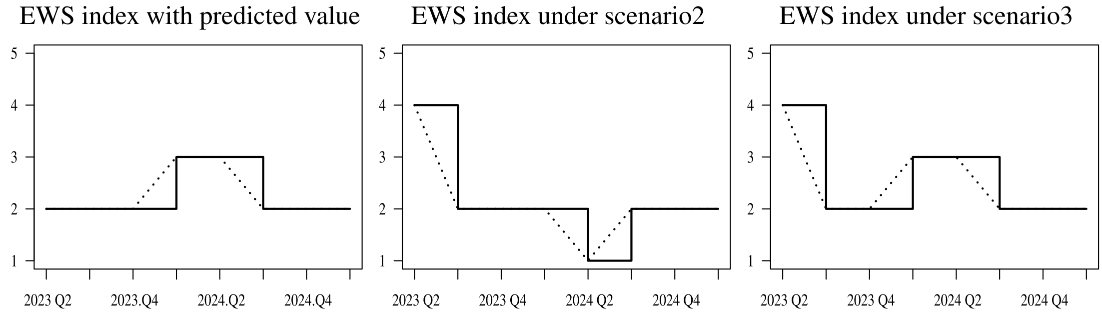

Fig. 10. Forecasts under the three scenarios. Notes: Each figure indicates the EWS index under scenario 1,2,3.

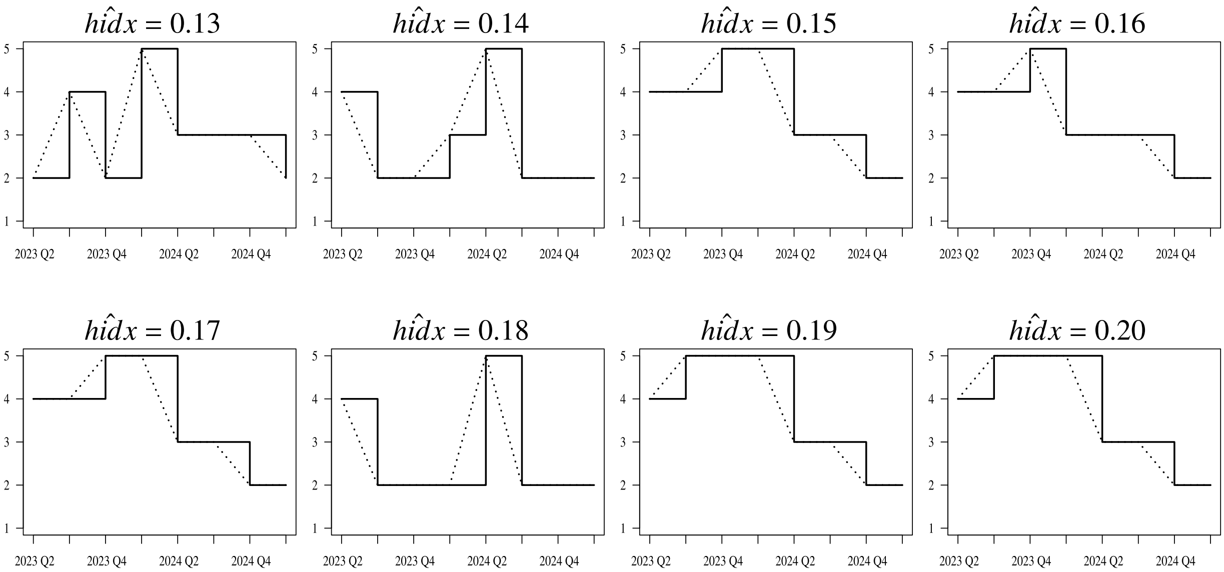

Fig. 11. Forecasts of EWS index with risk level 5.
Notes: $hi\hat{d}x$ denotes the growth rate of the house price index; in this study, we assume that the house price index grows at a constant growth rate for the next eight quarters.

level 4 in the first quarter of 2023, and for the next five quarters, it is almost fixed at risk level 2. The immediate response of the increasing growth rate of the house price index can be confirmed in the first quarter.

The last scenario assumes that the house price index rises by 1%p each quarter, which is a very strong assumption. Third figure of Fig. 10 shows the results of Scenario 3. The results are similar to scenario 2 in that the risk level reaches level 4 in the first quarter, after which the risk level fluctuates between levels 2 and 3. None of these scenarios exhibit a risk level of 5. However, we can see that house price index growth rate affects the time of the crisis. The higher the house price index growth rate, the faster the high risk level appears, and after a while, it goes back to lower risk level. Moreover, the risk levels with an increase house price index assumption exhibit higher value than the risk level that does not. Also, as we can see in Fig. 9, since there are no abnormalities in the predictions of other variables, it is acceptable that the scenario does not exhibit risk level of 5.

In order to find the house price index level at which the EWS model warns of level 5 risk, we arbitrarily change the growth rate of the house price index. Fig. 11 shows the predicted EWS index when the growth rate of the house price index is more than 13%. The trained model begins to warn of a risk level 5 once the house price index starts to grow at a 13% growth rate. This analysis show that the growth rate of house price index determines the time of crisis occurrence. The results show that the higher the rate of house price growth, the sooner the high risk level appears.

4. Concluding remarks

This study develops an early warning system for household debt risk, which has recently emerged as an important issue, and uses several macroeconomic variables and internet search data to predict the household debt risk. As the debt-to-GDP and debt-to-DPI ratios have gradually risen in recent years, accurate early warning of household debt risk can help adequately prepare for risks and establish appropriate economic policies. This study uses an index that is appropriately transformed with fractional differences in the debt-to-GDP ratio.

Several studies have analyzed early warning models for the financial crisis. These models are analyzed using binomial probability variables. However, in reality, a crisis does not occur suddenly. Therefore, in this study, the household debt crisis phases are divided into five levels with a moving average baseline.

Moreover, the previous models are in linear fashion, which oversimplifies the complex relationship between the dependent and independent variables. Of course, if there are clear and large numbers of input variables that can explain household debt risk, the predictive power of the linear model could be higher than that of other complex non-linear models. In reality, however, there is a limit to selecting a large number of clear input variables. Therefore, in this study, we use a deep learning model that does not require any function to set the threshold and bounds.

In addition, this study establishes a model using internet search data, which have been widely used recently. Many recent studies have reported statistically significant input variables. Consequently, the inclusion of Internet search data as input variables was intended to improve the predictive power of the EWS index. Based on the estimated results using the deep learning model, this study predicted the household debt risk level from the second quarter of 2021 to the first quarter of 2023. The analysis shows that there is no risk of household debt reaching level 5 in the forecast period. However, from the first quarter of 2022, the risk level reaches level 4, and 1st quarter of 2023, the forecast value returns to risk level 4. These repeated instances of high risk suggest that the government must continue to look carefully at trends in household debt and related variables. Moreover, by simulating the house price index, this study find that if the house price index growth rate exceeds 13%, the household debt EWS index may be able to exhibit risk level of 5.

The limitations of this study are as follows: First, during the COVID-19 pandemic, central banks in each country may have used accommodative monetary policy. Under these circumstances, the analysis of this study may have provided biased estimates. It can be a good study in the future if the analysis is conducted excluding the period of the pandemic or if an early warning model is developed that includes changes in the government's internal regulations. Second, we only have 85 observations for each variable. The advantage of forecasting with deep learning is its accurate predictive power with a large amount of data—whether this advantage remains when the data size is relatively small is still an open question. Third, it is hard to identify the effect of some specific factors in the DNN and LSTM models. This is due to the fact that considered factors affect the EWS index in a complex non-linear way. The above issues will be addressed in future research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the link to my data at the attached file

References

Abbas, O. M. (2005). Neural networks in business forecasting. *International Journal of Computer*, *19*(1), 114–128.

Abbas, O. M. (2017). A comparative study on business forecasting accuracy among neural networks and time series. *International Journal of Computer*, *26*(1), 175–183.

Alessi, L., & Detken, C. (2018). Identifying excessive credit growth and leverage. *Journal of Financial Stability*, *35*, 215–225, Network models, stress testing and other tools for financial stability monitoring and macroprudential policy design and implementation.

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

Arcand, J. L., Berkes, E., & Panizza, U. (2015). Too much finance?. *Journal of Economic Growth*, *20*(2), 105–148.

Barnes, S., & Young, G. (2003). The rise in US household debt: assessing its causes and sustainability. *Working paper*, Bank of England.

Berg, A., & Pattillo, C. (1999). What caused the asian crises: an early warning system approach. *Economic Notes*, *28*(3), 285–334.

Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, *73*(1), 151–184.

Caggiano, G., Calice, P., & Leonida, L. (2014). Early warning systems and systemic banking crises in low income countries: A multinomial logit approach. *Journal of Banking & Finance*, *47*, 258–269.

Caggiano, G., & Leonida, L. (2009). International output convergence: evidence from an autocorrelation function approach. *Journal of Applied Econometrics*, *24*(1), 139–162.

Cecchetti, S. G., Mohanty, M. S., & Zampolli, F. (2011). The real effects of debt. *BIS working paper*, Bank for International Settlements.

Cho, K., van Merrienboer, B., Bahdanau, D., & Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 conference on empirical methods in natural language processing* (pp. 1724–1734). Doha, Qatar.

Connor, J., Atlas, L. E., & Martin, D. R. (1992). Recurrent networks and NARMA modeling. In *Advances in neural information processing systems* (pp. 301–308).

Cynamon, B. Z., & Fazzari, S. M. (2008). Household debt in the consumer age: source of growth–risk of collapse. *Capitalism and Society*, *3*(2).

Dasgupta, S., & Osogami, T. (2017). Nonlinear dynamic Boltzmann machines for time-series prediction. In *Proceedings of the AAAI conference on artificial intelligence*.

Diaz, J., & Osler, T. (1974). Differences of fractional order. *Mathematics of Computation*, *28*(125), 185–202.

Duca, M. L., & Peltonen, T. A. (2013). Assessing systemic risks and predicting systemic events. *Journal of Banking & Finance*, *37*(7), 2183–2195.

Galil, K., & Soffer, G. (2011). Good news, bad news and rating announcements: An empirical investigation. *Journal of Banking & Finance*, *35*(11), 3101–3119.

Granger, C. W. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, *14*(2), 227–238.

Granger, C. W., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, *1*(1), 15–29.

Graves, A. (2012). Long short-term memory. In *Supervised sequence labelling with recurrent neural networks* (pp. 37–45). Berlin: Springer.

Haslett, J., & Raftery, A. E. (1989). Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource. *Journal of the Royal Statistical Society. Series C. Applied Statistics*, *38*(1), 1–21.
Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, *18*(7), 1527–1554.
Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.
Holopainen, M., & Sarlin, P. (2017). Toward robust early-warning models: A horse race, ensembles and model uncertainty. *Quantitative Finance*, *17*(12), 1933–1963.
Hosking, J. (1981). Equivalent forms of the multivariate portmanteau statistic. *Journal of the Royal Statistical Society. Series B.*, *43*(2), 261–262.
Jain, A., & Kumar, A. M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, *7*(2), 585–592.
Kamin, S. B., & Babson, O. D. (1999). The contributions of domestic and external factors to Latin American devaluation crises: an early warning systems approach. *no.645*, In *International finance discussion papers, the federal reserve board of governors*. Board of Governors of the Federal Reserve System.
Kamin, S. B., Schindler, J. W., & Samuel, S. L. (2001). The contribution of domestic and external factors to emerging market devaluation crises: An early warning systems approach. *no.711*, In *International finance discussion papers*. Board of Governors of the Federal Reserve System.
Kaminsky, G., Lizondo, S., & Reinhart, C. M. (1998). Leading indicators of currency crises. *IMF Economic Review*, *45*(1), 1–48.
Kang, J. K. (2017). The effects of household debt on consumption and economic growth. In *BOK economic research*. Bank of Korea.
Karasulu, M. (2008). Stress testing household debt in Korea. In *IMF working paper*. International Monetary Fund.
Kim, H. J., Lee, D., Son, J. C., & Son, M. K. (2014). Household indebtedness in Korea: Its causes and sustainability. *Japan and the World Economy*, *29*, 59–76.
Kim, H. J., Son, J. C., & Yie, M.-S. (2017). House price dynamics with household debt: the korean case. *Asian Economic Journal*, *31*(1), 39–59.
Malagrino, L. S., Roman, N. T., & Monteiro, A. M. (2018). Forecasting stock market index daily direction: a bayesian network approach. *Expert Systems with Applications*, *105*, 11–22.
Maxwell, A., Li, R., Yang, B., Weng, H., Ou, A., Hong, H., Zhou, Z., Gong, P., & Zhang, C. (2017). Deep learning architectures for multi-label classification of intelligent health risk prediction. *BMC Bioinformatics*, *18*(14), 121–131.
McLeod, A. I., & Hipel, K. W. (1978). Preservation of the rescaled adjusted range: A reassessment of the hurst phenomenon. *Water Resources Research*, *14*(3), 491–508.
Meniago, C., Mukuddem-Petersen, J., Petersen, M. A., & Mongale, I. P. (2013). What causes household debt to increase in South Africa? *Economic Modelling*, *33*, 482–492.
Michelacci, C., & Zaffaroni, P. (2000). (Fractional) beta convergence. *Journal of Monetary Economics*, *45*(1), 129–153.
Niesert, R. F., Oorschot, J. A., Veldhuisen, C. P., Brons, K., & Lange, R.-J. (2020). Can google search data help predict macroeconomic series? *International Journal of Forecasting*, *36*(3), 1163–1172.
Nyman, R., Kapadia, S., & Tuckett, D. (2021). News and narratives in financial systems: exploiting big data for systemic risk assessment. *Journal of Economic Dynamics & Control*, *127*(104119).
Park, S. Y. (2017). A study on the development of early warning systme for household debt risk using big data. In *BOK report paper*. Bank of Korea.
Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural networks. In *International conference on machine learning* (pp. 1310–1318).
Qin, X., & Luo, C. (2014). Capital account openness and early warning system for banking crises in G20 countries. *Economic Modelling*, *39*, 190–194.
Rambaccussing, D., & Kwiatkowski, A. (2020). Forecasting with news sentiment: Evidence with UK newspapers. *International Journal of Forecasting*, *36*(4), 1501–1516.
Rather, A. M., Agarwal, A., & Sastry, V. (2015). Recurrent neural network and a hybrid model for prediction of stock returns. *Expert Systems with Applications*, *42*(6), 3234–3241.
Samko, S. G., & Ross, B. (1993). Integration and differentiation to a variable fractional order. *Integral Transforms and Special Functions*, *1*(4), 277–300.
Scherer, R., Kalla, S. L., Tang, Y., & Huang, J. (2011). The Grünwald–Letnikov method for fractional differential equations. *Computers & Mathematics with Applications*, *62*(3), 902–917.
Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series. In *2018 17th IEEE international conference on machine learning and applications* (pp. 1394–1401).
Smales, L. A. (2016). News sentiment and bank credit risk. *Journal of Empirical Finance*, *38*, 37–61.
Song, W. (2010). Building an early warning system for crude oil price using neural network. *East Asian Economic Review*, *14*(2), 79–109.
Stockhammer, E., & Wildauer, R. (2018). Expenditure cascades, low interest rates or property booms? determinants of household debt in OECD countries. *Review of Behavioral Economics*, *5*(2), 85–206.
Tanaka, K., Kinkyo, T., & Hamori, S. (2016). Random forests-based early warning system for bank failures. *Economics Letters*, *148*, 118–121.
Ticknor, J. L. (2013). A bayesian regularized artificial neural network for stock market forecasting. *Expert systems with applications*, *40*(14), 5501–5506.
Tsai, F.-T., Lu, H.-M., & Hung, M.-W. (2016). The impact of news articles and corporate disclosure on credit risk valuation. *Journal of Banking & Finance*, *68*, 100–116.
Zabai, A. (2017). Household debt: recent developments and challenges. *BIS Quarterly Review*.
Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159–175.
Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting*, *14*(1), 35–62.
Zhang, Z., & Sabuncu, M. R. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. In *Proceedings of the 32nd international conference on neural infornation processing systems* (pp. 8792–8802).