Quantitative Finance and Economics

http://www.aimspress.com/journal/QFE

QFE, 1(3): 219–252
DOI:10.3934/QFE.2017.3.219
Received date: 1 May 2017
Accepted date: 26 July 2017
Published date: 12 October 2017

Research article

Lending Sociodynamics and Drivers of the Financial Business Cycle

Raymond J. Hawkins [1,2,*] **and Hengyu Kuang** [1]

[1] Economics Department, University of California Berkeley, Berkeley, CA 94720, USA
[2] College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA

* **Correspondence:** Email: hawkins@econ.berkeley.edu; Tel: +1-415-254-0788.

Abstract: We extend sociodynamic modeling of the financial business cycle to the Euro Area and Japan. Using an opinion-formation model and machine learning techniques we find stable model estimation of the financial business cycle using central bank lending surveys and a few selected macroeconomic variables. We find that banks have asymmetric response to good and bad economic information, and that banks adapt to their peers' opinions when changing lending policies.

Keywords: economic instability; Financial Instability hypothesis; Sociodynamics; opinion formation; systemic risk; Random Forest; Fokker-Planck equation

1. Introduction

In Post-Keynesian economic thought money is endogenously determined by economic activity in a modern economy: "loans create deposits" and "deposits generate reserves." New money is created by lending and is destroyed when loans are paid back. Banks act not as intermediaries bound by reserves that lend deposits; rather, they are central economic actors that create money in the forms of loans (Werner, 2016). Thus, an understanding of the drivers and dynamics of lending behavior is central to an understanding of the financial business cycle.

Central bankers have long recognized that banks create money. The Central Bank of Canada governor, Towers (1939), observed that "[e]ach and every time a bank makes a loan, new bank credit is created - new deposits - brand new money." The New York Federal Reserve Vice President, Holmes (1969), recognized the Fed's operational constraints on the money supply: "In the real world, banks extend credit, creating deposits in the process, and look for the reserve later." McLeay et al. (2014) of the Bank of England clarified misconceptions concerning money creation a modern economy thusly[*]: "banks do not act simply as intermediaries, lending out deposits that savers place with them, and nor

[*]For an empirical demonstration of money creation by a commercial bank from loan creation to money transfer see Werner (2014, 2016).

do they ‘multiply up’ central bank money to create new loans and deposits.”

This perspective is also central to the Financial Instability hypothesis (Minsky, 1977, 2008) in which Minsky integrated Keynes’ observation concerning the social psychology of bankers (Keynes, 1931)

> “[a] ‘sound banker, alas! is not one who foresees danger and avoids it, but one who, when he is ruined, is ruined in a conventional way along with his fellows, so that no one can really blame him.”

with Keynes’ view of the importance of cognitive and social psychology as a drivers of fluctuations in the financial and real economy (Keynes, 1937):

> “We assume that the present is a much more serviceable guide to the future than a candid examination of past experience would show it to have been hitherto. … We assume that the existing state of opinion as expressed in prices and the character of existing output is based on a correct summing up of future prospects, … Knowing that our own individual judgement is worthless, we endeavor to fall back on the judgement of the rest of the world which is perhaps better informed. That is, we endeavor to conform with the behavior of the majority or the average. … Now a practical theory of the future based on these three principles has certain marked characteristics. In particular, being based on so flimsy a foundation, it is subject to sudden and violent changes.”

Lenders interact within a social structure that has norms for group members’ common behaviors and network mechanisms controlling the flow of information. They rely on other lenders to justify their economic actions and this gives rise to the phenomenon of “social embeddedness” in banking (Granovetter, 2005).

In this paper we provide further evidence of the importance of social psychology in the generation of the financial business cycle in a capitalist economy by modeling the evolution of lending in terms of psychological and economic drivers that matter to bankers. Our analysis employs the opinion-formation model introduced by Weidlich (1972) and later elaborated by Weidlich and Hagg (1983) with which the social psychology of economic agents can be included in the dynamics of economic incentives.[†] Our analysis develops earlier work on the sociodynamics of lending (Hawkins, 2011) through the use of simulation developed by Lux (1997, 2009a,b) and recently applied to the problem of lending opinion formation by Ghonghadze and Lux (2012, 2016). The work of Ghonghadze and Lux (2016) is of particular interest as it demonstrates that a rich set of lending-opinion data – the Federal Reserve’s Senior Loan Officer Opinion Survey on Bank Lending Practices (SLOOS) – can be successfully incorporated into the opinion-formation model. Our interest is in extending this work beyond the United States to economies where lending-opinion data sets may be less complete. The results of Ghonghadze and Lux (2016) suggests that machine-learning approaches may be of use in identifying the subset of lending-opinion data that is key to lending-opinion formation and a goal of our paper is to demonstrate that this is so.

To this end after an introduction of the opinion-formation model in Section 2 we begin our analysis in Section 3.1 with a demonstration that the U.S. lending study of Ghonghadze and Lux (2016) can be reproduced, both to validate their (and our) work with this data set and to motivate the use of random forests – a machine-learning approach to variable reduction – for this problem. We examine the use

[†] See also Weidlich (2000) and Helbing (2010).

of random forests on the U.S. lending data set and then apply it to lending surveys provided by the European Central Bank, and the Bank of Japan in Sections 3.2 and 3.3, respectively. We close with a reflection on the results, both expected and unexpected, of our study in Section 4.

2. The Opinion-Formation Model

The competitiveness of a loan product is determined by the credit spread – the difference between the interest rate of a loan product and the bank's cost of funds – that a bank charges. Setting the level of the spread is a key task of banking and the factors that bear on this key decision are the drivers of the opinion-formation model. As in Ghonghadze and Lux (2016) we consider a collection of $2N$ bankers facing this question at a point in time t. Of these, n_t^+ are positively disposed to lend and n_t^- are against lending at that time; by construction $n_t^+ + n_t^- = 2N$. This state of opinion can be represented by the difference:

$$2n_t = n_t^+ - n_t^- \tag{1}$$

where $n_t \in [-N, N]$. When $n_t = N$, all bankers choose to lend, and when $n_t = -N$ all choose not to. The lending-opinion index, or the average lending sentiment, at time t can be written as the normalized difference of opinion:

$$x_t := \frac{n_t}{N} = \frac{n_t^+ - n_t^-}{2N}. \tag{2}$$

which, as noted by Ghonghadze and Lux (2016), is precisely how empirical diffusion indices are calculated.

One also assumes there to be a lending-configuration dependent individual lender transition probability per unit time representing the change from being willing to lend to being unwilling to lend which we write as $p_{+-}(n)$; and similarly for the transition from being unwilling to lend to being willing to lend $p_{-+}(n)$. Finally, if one considers the simplest case where the lending configuration changes by the change of lending opinion of only a single lender per unit time, the equation of motion for the probability that the lending community is in configuration n at time t, $p(n;t)$, is given by the master equation

$$\begin{aligned}\frac{\partial p(n;t)}{\partial t} &= \left[w_\downarrow(n+1)p(n+1;t) - w_\downarrow(n)p(n;t)\right] \\ &\quad + \left[w_\uparrow(n-1)p(n-1;t) - w_\uparrow(n)p(n;t)\right]\end{aligned} \tag{3}$$

where $w_\uparrow(n) = (N-n)p_{+-}(n)$ and $w_\downarrow(n) = (N+n)p_{-+}(n)$. The terms in the first set of brackets on the right-hand side of Eq. (3) indicate the change to $p(n;t)$ resulting from a single lender moving from the position of being willing to lend to being unwilling to lend $w_\downarrow(n)$: the first term in this bracket describes the increase in configuration probability due to transitions from the configuration with one more net lender while the second term describes the decrease in configuration probability from transitions to the configuration with one less net lender. Similarly, terms in the second set of brackets on the right-hand side of Eq. (3) indicate the change to $p(n;t)$ that result from a single lender moving from the position of being unwilling to lend to being willing to lend $w_\uparrow(n)$: the first term in this bracket describes the increase in configuration probability due to transitions from that with one less net lender while the second term describes the decrease in configuration probability from transitions to that with one more net lender.

This master equation description of probability dynamics is, for our purposes, more conveniently handled using the related FokkerPlanck equation

$$\frac{\partial P(x;t)}{\partial t} = -\frac{\partial}{\partial x}[K(x)P(x;t)] + \frac{1}{2N}\frac{\partial^2}{\partial x^2}[Q(x)P(x;t)] \tag{4}$$

which describes the evolution of the probability of the lending sentiment $P(x;t) = Np(Nx;t)$, and where $K(x) = (w_\uparrow(n) - w_\downarrow(n))/N$ is the drift coefficient and $Q(x) = (w_\uparrow(n) + w_\downarrow(n))/N$ is the diffusion coefficient.[‡]

One can link this model to the view of bank decision making advanced by Keynes (1931) and (Minsky, 1977, 2008) by summarizing that view as follows: (i) banks make their decisions in part by relying on their own independent observations and analysis, but (ii) bankers are willing to adapt to the peers opinion once it becomes the majority, and that (iii) bankers' preference and their willingness to adapt may vary over time. This summary can be incorporated into the opinion-formation model by writing the transition probabilities as[§]

$$w_\uparrow(x) = \frac{n^-}{2N}\nu \exp(U(\cdot)) = (1-x)\nu \exp(U(\cdot)) \tag{5}$$

$$w_\downarrow(x) = \frac{n^+}{2N}\nu \exp(-U(\cdot)) = (1+x)\nu \exp(-U(\cdot)) \tag{6}$$

and

$$U(x_t, Z_t; \theta) = \alpha_0 + \alpha_1 x_t + \sum_{i=2}^{M} \beta_i Z_i \tag{7}$$

where $n^+/2N$ and $n^-/2N$ measure the attitude towards lending, ν is the speed with which lending opinion changes, α_0 measures the individual component of lending decisions, α_1 measures the influence on lending decision of other lenders, the Z_i are relevant economic data such as GDP and unemployment that affect lending decisions and which bear on the lending decision with weight $\beta_i, i = 1, \ldots, M$, and $\theta = \{\nu, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$ denotes the collection of coefficients to be estimated.

Our simulation implementation follows that of Ghonghadze and Lux (2016) by writing the drift and diffusion terms as

$$K(x_t, Z_t; \theta) = 2\nu \cosh(U(x_t, Z_t; \theta))(\tanh(U(x_t, Z_t; \theta)) - x_t) \tag{8}$$

and

$$Q(x_t, Z_t; \theta) = 2\nu \cosh(U(x_t, Z_t; \theta))(1 - x_t \tanh(U(x_t, Z_t; \theta)))/N \tag{9}$$

and numerically evaluating the stochastic differential equation

$$dx_t = K(x_t, Z_t; \theta)dt + \sqrt{Q(x_t, Z_t; \theta)}dW_t \tag{10}$$

associated with our Fokker-Planck equation where W_t is the standard Wiener process.

[‡]The term $K(x)$ represents a force that acts on opinion formation which, given the dependence on α_1 through Eqs. (5) – (7) links this force to the opinion of others, has a social component. This force implies the existence of a potential $V(x)$ where $K(x) = -\partial_x V(x)$ which likewise has a social component (Helbing, 2010). The social component of the force represents psychic tension which a banker will seek to eliminate by behavior that moves them to a minimum of the social field $V(x)$. At the minimum psychic tension vanishes; eliminated in this case by conforming to lending opinion of others.

[§]See Weidlich and Haag (1983), Lux (1997, 2009a,b) and Ghonghadze and Lux (2016).

Empirical application with this model requires estimating the parameters θ. Ghonghadze and Lux (2012) solved Eq. (4) using the Crank-Nicolson method together with the log Maximum Likelihood on the solution $P(x;t)$ to estimate θ. This method, while rigorous, is computationally expensive. Ghonghadze and Lux (2016) later introduced the use of the more efficient Quasi Maximum Likelihood Estimate (QLME) method where P is approximated as a conditional normal distribution and:

$$E(x_{t+1}|x_t) = x_t + K(x_t, Z_t; \theta)\Delta t \tag{11}$$

$$Var(x_{t+1}|x_t) = Q(x_t, Z_t; \theta)\Delta t. \tag{12}$$

Given T observations $\{x_1, \cdots, x_T\}$, θ is estimated by maximizing the log-likelihood function

$$L = \sum_{t=1}^{T-1} \ln(N(x_{t+1}|x_t, Z_t; \theta)), \tag{13}$$

where $N(x_{t+1}|x_t, Z_t; \theta)$ is normal density with mean $E(x_{t+1}|x_t)$ and variance $Var(x_{t+1}|x_t)$. We solve L numerically using the Newton Conjugate Gradient method in Scipy's Generic likelihood model.

In applications of this model one often encounters references to different versions of the model based on the number of terms on the right-hand side of Eq. (7) that are retained; calibration with all $Z_i = 0$ being a "model" and those with differing numbers of Z_i being referred to as other "models" (see, for example, Ghonghadze and Lux (2016) and references therein). Operationally our references in this manner later in this paper are to the opinion-formation model, Eqs. (8)–(13), with similarly different expressions of Eq. (7).

3. Credit Dynamics

Our datasets of interest are central bank lending surveys. The Federal Reserve (Fed), the European Central Bank (ECB), and the Bank of Japan (BOJ) all publish quarterly surveys of their bank's lending practices.[¶] The Fed survey is the oldest, dating back to 1990. Japan followed suit a decade later and the European survey launched in 2003 with details on each country within the Union. All surveys ask banks whether they change credit standards and spreads for firms and households, and why they change their lending policies.

We begin our study with an examination of the Fed survey. Working with the Fed data allows us to validate our approach against the earlier work of Ghonghadze and Lux (2016) and to provide context for our introduction of a machine-learning technique (random forests) into the variable selection process. We then extend this analysis to the surveys of the ECB and the BOJ.

3.1 Credit Dynamics in the United States

The Fed's Senior Loan Officer Opinion Survey on Bank Lending Practices (SLOOS) is conducted at the beginning of each calendar quarter, covering up to 60 large domestic commercial banks and up to 24 large foreign banks. All have asset greater than \$3 billion or at most 5% of commercial and

[¶]The Bank of England also began publishing a survey after the beginning of the Great Recession. Unfortunately, their survey data is weighted by the level of importance, and is therefore incompatible with our model assumption. Since the raw data is not yet available the application of our approach to British banking sentiment cannot proceed at this time.

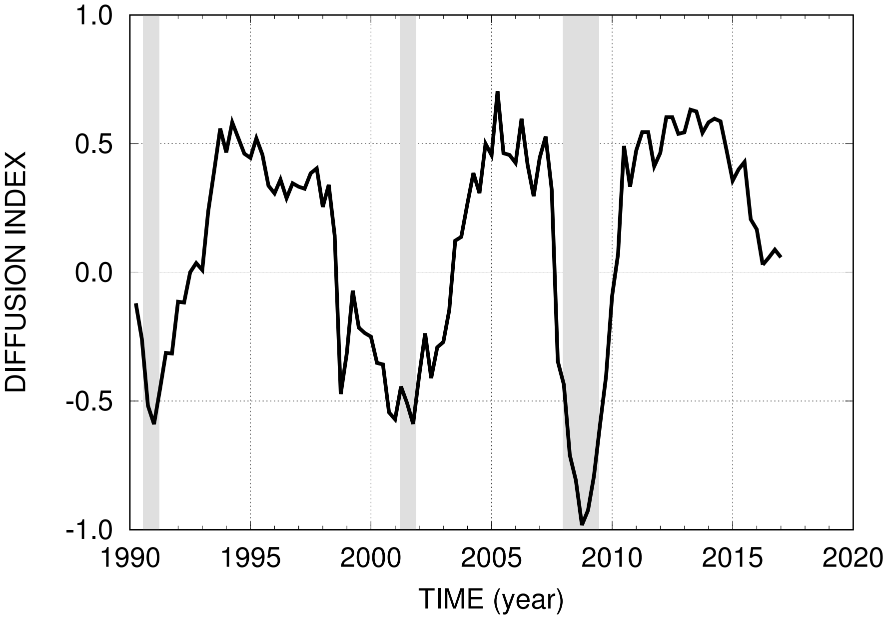

Figure 1. The credit spread diffusion index (DICS) for the United States. NBER recessions are indicated by the grey bars.

industrial loans to total assets. We focus on question 2(d) of the survey which concerns the credit spread on commercial and industrial loans for middle and large firms:

> *2. For applications for C&I loans or credit lines – other than those to be used to finance mergers and acquisitions – from large and middle-market firms and from small firms that your bank currently is willing to approve, how have the terms of those loans* ***changed over the past three months****? (emphasis added) (d) Spread of loan rates over your bank's cost of funds (wider spreads = tightened, narrower spread = eased)*

Five options are available: tightened considerably, tightened somewhat, remained basically unchanged, eased somewhat, and eased considerably. On average around 70 banks answer this question. From the responses to this question a diffusion index (DI) can be constructed as the the net percent of banks that eased less those that tightened, and the results of such a construction are shown in Fig. 1.[‖] The dynamic range and speed of change of the DICS in the U.S. is striking and reveals, interestingly, that banks collectively tighten credit well before the last two recessions and that they loosen credit quickly once a recession has ended.

The survey gives a series of factors for banks to choose if they changed lending policies in the past quarter. The following six factors have been consistently reported since 1999:

> *If your bank has tightened or eased its credit standards or its terms for C&I loans or credit lines over the past three months, how important have been the following possible reasons for the change? (Please respond to either A, B, or both as appropriate and rate each possible reason using the following scale: 1 = not important, 2 = somewhat important, 3 = very*

[‖]In the Fed's report, this diffusion index is the net percentage of banks that tightened less those that eased.

important)

(a) *Deterioration (improvement) in current or expected capital position.*

(b) *Less (more) favorable or more (less) uncertain economic outlook.*

(c) *Worsening (improvement) of industry-specific problems.*

(d) *Less (more) aggressive competition from others.*[**]

(e) *Reduced (increased) risk tolerance for risk.*

(f) *Decreased (increased) liquidity in the secondary market for these loans.*

The data associated with these questions are only available in the pdf version of the Fed reports. Thus to calculate an associated sentiment concerning lending we extracted the data from the reports for the period 1999–2017 and recalculated the mean for each reason.[††] The means of these reasons – potential drivers of lending sentiment and the DICS from Fig. 1 – are shown together in Fig. 2. Beginning with the top-left panel we see that negative values about uncertainty correlates well with the DICS; positive values less so. The opposite is seen concerning competition in the upper-right panel where a positive view shows the greatest dynamic range. A bank's capital position is remarkably uncorrelated with the DICS as shown in the left-hand middle panel, although negative industry-specific issues correlate strongly with the the DICS. The panels in the bottom row of Fig 2 show that negative risk tolerance and liquidity have greater correlation with DICS than do their positive counterparts.

There is also an interesting asymmetry in the two largest potential drivers of the DICS change. When banks tighten lending uncertainty is the main driving force, but uncersainty has almost no effect in easing credit. By contrast, competition is the main driving force when banks ease lending. Such repeated asymmetric response during business cycles provide an empirical support for the assumptions of the opinion-formation model. Banks formulate their economic outlook, but they are also impacted by their peers' decisions. When they are not sure about their customers' creditworthiness they tighten lending. But once their peers "control the Street" (competition increases) the state of the economy – and, by implication, the customer – becomes less important in driving credit expansion and lending decisions. Curiously, lending decisions found to be much less dependent on capital positions and loan liquidity. Finally we note that lending decision are less dependent on capital position and loan liquidity. The former is consistent with the post-Kenyesian interpretation of endogenous money and the fact that at least some problems with capital structure can be remedied through short-term borrowing from the central bank of other commercial banks. The latter suggests a lack of existential dependence for most banks on loan trading.

The high rate of change – both positive and negative – of the DICS and associated drivers (e.g., uncertainty and competition) during the 2008 Great Recession suggest that opinion changed in a manner known as a *phase transition*; a phenomena that can be simulated using the opinion-formation model (Weidlich and Haag, 1983). This has important implications for systemic risk management since the onset of a phase transition is less about the event that caused it (which has historically been

[**]The surveys before 2001 Q3 split this reason into 2 parts: competition from nonbank lenders and other banks. We take the average of their value.

[††]Responses were weighted by the scale of importance (0, 1, 2) as well as their fraction of the total number of banks responding. The reasons contributing to tighten lending policies are set to be negative. If a reason has value 0, it means either (i) no bank eased or tightened lending or (ii) all banks think the reason is not important. If 2 (-2), all banks consider it is very important; 1 (-1) means banks think it is somewhat important. Values in between mean some banks eased and/or tightened lending and have mixed opinions.

Table 1. Estimated coefficients for the minimal opinion-formation models for the United States.

Parameter	*Estimate*	*Standard Error*	*z*	*P > |z|*	*95% C.I.*
Model 1: Eq. (14)					
ν	1.6657	0.204	8.160	0.000	1.266 2.066
α_0	0.0009	0.020	0.047	0.963	-0.038 0.040
α_1	1.0440	0.049	21.126	0.000	0.947 1.141
$\mathcal{LL}$	44.74				
AIC	-85.74				
BIC	-80.13				
Model 2: Eq. (15)					
ν	1.6767	0.207	8.101	0.000	1.271 2.082
α_0	0.0651	0.044	1.485	0.138	-0.021 0.151
α_1	0.8793	0.113	7.810	0.000	0.659 1.100
α_2	1.2276	0.120	10.191	0.000	0.991 1.464
$\mathcal{LL}$	46.10				
AIC	-88.20				
BIC	-82.86				

Note: The models were estimated using 107 observations from 70 banks. $\mathcal{LL}$, AIC, and BIC are log-likelihood, Akaike information criteria, and Bayesian information criteria, respectively.

the focus of risk management) and more about the state of the system (economy) at the point when the phase transition occurs.

Next we turn to the opinion-formation model and some of the specifications examined by Ghonghadze and Lux (2016). When there are no exogenous effects Z_i and banks only look to their peers for guidance on lending the dynamics driver U can be written

$$U(x_t, Z_t; \theta) = \alpha_0 + \alpha_1 x_t \tag{14}$$

where $\theta = \{\nu, \alpha_0, \alpha_1\}$ or

$$U(x_t, Z_t; \theta) = \alpha_0 + \alpha_1 x_t^+ + \alpha_2 x_t^- \tag{15}$$

where

$$\theta = \{\nu, \alpha_0, \alpha_1, \alpha_2\} \ , \quad x_t^+ = \begin{cases} x_t & x_t > 0 \\ 0 & x_t \leq 0 \end{cases} \ , \quad \text{and} \quad x_t^- = \begin{cases} x_t & x_t \leq 0 \\ 0 & x_t > 0 \end{cases} \tag{16}$$

Ghonghadze and Lux (2016) refer to the specification given by Eq. (14) as "Model 1" and that given by Eq. (15) as "Model 2" and we shall do so as well. Model 1 assumes that contemporaneous lending opinion is the sole driver and that positive and negative values of lending sentiment have the same weight. Model 2 relaxes the assumption of identical weight by allowing for a difference in how positive and negative values of lending sentiment impact the lending decision.

To ensure that our approach was consistent with that of Ghonghadze and Lux (2016) we estimated these specifications, the results of which are shown in Table 1. Model 1 indicates the current lending opinion, α_1, has significant impact and the opinion turn-over rate, ν, suggests banks change their credit opinions quickly. Model 2 yields coefficients for positive (α_1) and negative (α_1) lending opinion that are slightly different, but they overlap within two-standard-deviation region. Comparing Models 1 and 2 on an information basis one would prefer Model 2 slightly as the Akaike Information Criterion (AIC) and Bayesian Information Criterion(BIC) for this model are somewhat better than those of Model 1.

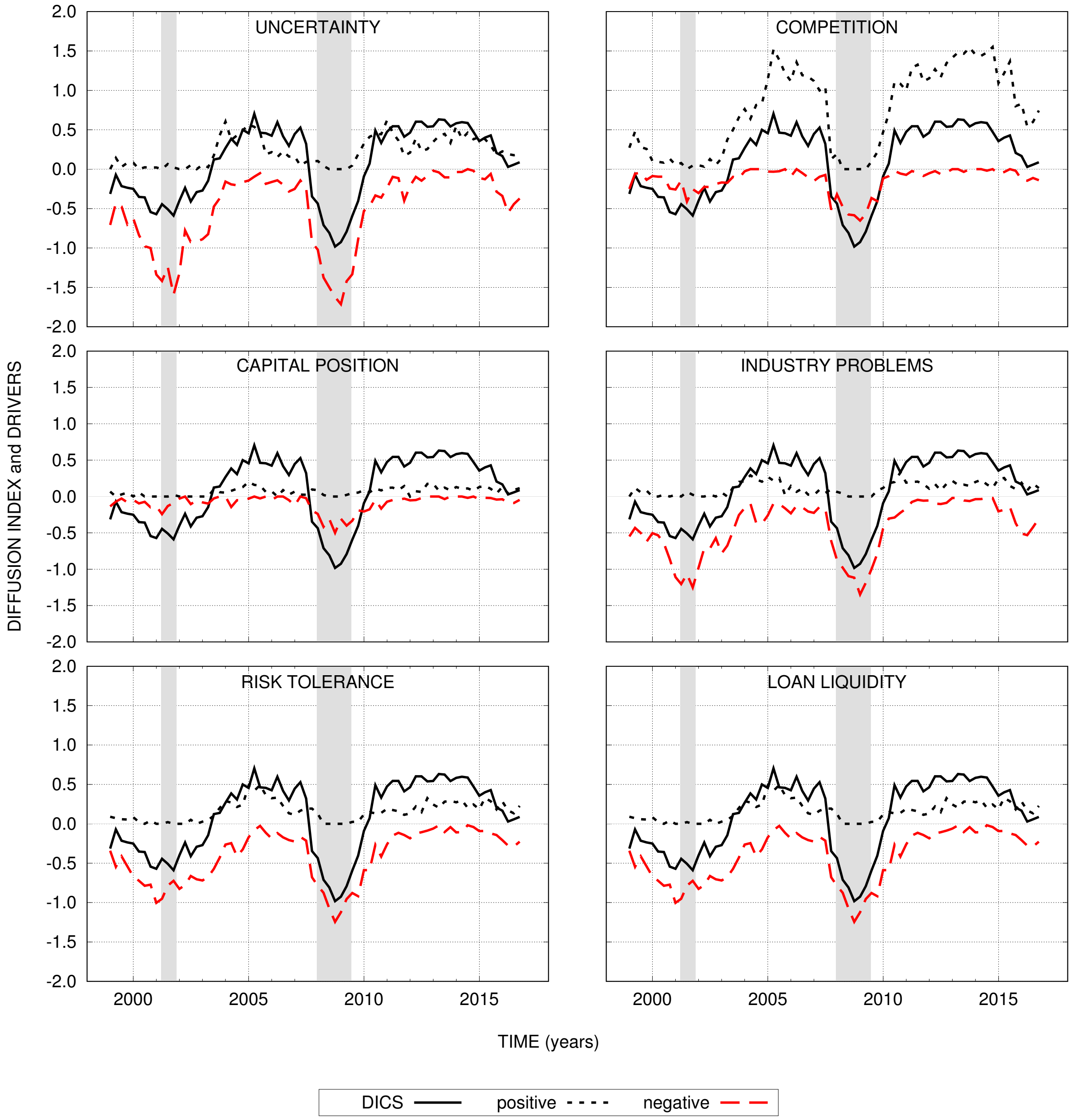

Figure 2. The credit spread diffusion index together with potential drivers of the diffusion index in the United States.

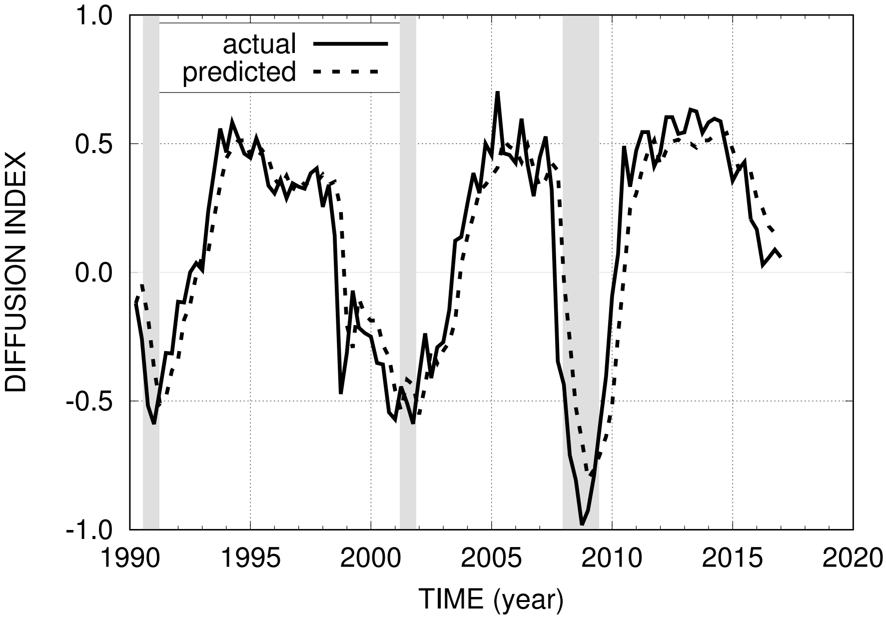

Figure 3. The observed credit spread diffusion index for the United States together with the simulation result using opinion-formation model given by Eq. (15).

A simulation of the DICS using Model 2 is shown in Fig. 3. The predicted values follow the actual data well. In this simulation the simulation the credit spread variables (x_{t+}, x_{t-}) in Eq. (15) are updated using the survey data rather than the simulated values. This level of agreement between the simulated and observed DICS suggests that the current opinion index of credit spread has strong predictive power concerning the credit dynamics for the next quarter.

Examining the stability of the Model 2 coefficients yielded two interesting results. First, the coefficients are stable over time, meaning banks are constantly looking at each other, and splitting the diffusion index may not be able to capture banks' asymmetric reaction to their peers. Second, the model estimation becomes stable after 60 data points. The model is also time independent once it is conditioned on the current quarter because the coefficients vary little after the data is randomly shuffled. These results and the conjuncture that only the current opinions have a much stronger impact on lending decisions suggest that most existing machine-learning algorithms are worth considering in the analysis of this diffusion index.

While the simulation results based on the current value(s) of the DICS are promising, the assumption that a banker would solely focus on current data is rather strong. This assumption was relaxed by Ghonghadze and Lux (2016) in their Model 3 where Eq. (15) is extended to include the past quarter's credit spread:

$$U(x_t, Z_t; \theta) = \alpha_0 + \alpha_1 x_t^+ + \alpha_2 x_t^- + \alpha_3 x_{t-1}^+ + \alpha_4 x_{t-1}^-, \tag{17}$$

where

$$\theta = \{\alpha_0, \alpha_1, \alpha_2\}, \qquad x_s^+ = \begin{cases} x_s & x_s > 0 \\ 0 & x_s \leq 0 \end{cases} \qquad x_S^- = \begin{cases} x_s & x_s \leq 0 \\ 0 & x_s > 0 \end{cases} \tag{18}$$

Table 2. Estimated coefficients for the opinion-formation model for the United States given by Eq. (17).

Parameter	*Estimate*	*Standard Error*	*z*	*P > |z|*	*95% C.I.*
ν	0.16088	0.201	8.015	0.000	1.215 2.002
α_0	0.0597	0.048	1.243	0.214	-0.034 0.154
α_1	1.0599	0.206	5.157	0.000	0.657 1.463
α_2	1.5344	0.206	7.463	0.000	1.131 1.937
α_3	-0.1808	0.200	-0.906	0.365	-0.572 0.211
α_4	-0.3545	0.207	-1.715	0.086	-0.760 0.051
$\mathcal{LL}$	48.22				
AIC	-86.43				
BIC	-73.11				

Note: The model was estimated using 107 observations from 70 banks. $\mathcal{LL}$, AIC, and BIC are log-likelihood, Akaike information criteria, and Bayesian information criteria, respectively.

The estimated coefficients for this model are shown in Table 2 where we see that the sum of coefficients for both positive and negative credit spread are close to those of Eq. (15). This suggests that banks' decision strongly rely on the lending opinion of the current period. Statistically, however there are issues with Model 3 that might recommend against it. Specifically, it is worse in terms of BIC and the coefficients of current and past DICS coefficients have conflicting signs; an indication of over-fitting.

Our examination of Models 1-3 finds that the results of Ghonghadze and Lux (2016) can be reproduced with a different – albeit related – method for solving Eq. (10) and validates our implementation of their model. On a shared dataset our parameters agreed to at least 4 decimal points and our variances had at most a 7% difference; excellent agreement given our different methodologies.

Next we include exogenous variables. Candidate variables selected using the SLOOS reasons for changing lending policy, the work of (Ghonghadze and Lux, 2016), and empirical research on the ECB lending surveys (Köhler-Ulbrich et al., 2016; Altavilla et al., 2015) are shown in Table 3. The diffusion index of loan demand (DILD) was built from question 4A of the SLOOS which asks how the C&I loan demand from large and middle market size firms changed over the past quarter. Excessive Bond Premium (EBP) measures the investors' risk appetite in the corporate bond market, has been shown to be a statistically significant leading indicator for recessions and statistically significant for credit supply.[‡‡] In this paper we use the EBP to measure the aggregate risk preference. The NASDAQ is the well-known stock index and the VIX is the volatility index known also as the "fear index". The non-performing loan (NPL) rate measures the performance of bank loans. Loan defaults take time to be observed, so 3 year and 5 year average cumulative default rate are key indicators for S&P credit ratings (Hawkins, 2011). In addition, Bouwman and Malmendier (2015) find that a history of under-capitalization impacts a bank's risk preference, and in this analysis 1-6 years are statistically significant. Real GDP (RGDP) is lagged two quarters because of its release data (e.g., first quarter's GDP is only available by the end of second quarter). The mean responses from the Survey of Professional Forecasters (SPF) were used to approximate a banks' estimate for the future quarter.

[‡‡]See Favara et al. (2016) & Gilchrist and Zakraj (2012), Favara et al. (2016), and Altavilla et al. (2015), respectively. This data is available on the FRED notes page. It assumes that the spread contains information on the expected default risk and independent risk preference. It is constructed by a simple unweighted average of credit spreads in corporate bonds from which one then subtracts each bonds' expected default risk implied in the spread through linear regression.

Table 3. Candidate variables for the opinion-formation model in the United States given by Eq. (7).

Variable	*Description*	*Release Date*	*S.A.*
DICS	Diffusion index of credit spread. Large & middle firms.	Mid next quarter	N
DILD	Diffusion index of loan demand. Large & middle firms.	Mid next quarter	N
EBP	Excessive bond premium. Quarter average, change.	Monthly	Y
NASDAQ	NASDAQ index. Quarter average, percent change.	Daily	N
VIX	CBOE volatility index. Quarter average, change.	Daily	N
NPL	Non performing commercial loans rate. De-meaned.	Mid next quarter	N
RGDP	RGDP growth rate. De-meaned.	End of next quarter	Y
SPF RGDP	SPF RGDP growth rate. De-meaned.	Mid current quarter	Y
Co. Profit	Corporate profit growth after tax. De-meaned.	End of next quarter	Y
SPF Co. Profit	SPF Co. profit growth De-meaned.	Mid current quarter	Y
Unemp	Unemployment rate U3, change . De-meaned.	Monthly	Y
SPF Unemp	SPF unemp rate, change. De-meaned.	Mid current quarter	Y
CPI	All items, change.	Monthly	Y
SPF CPI	SPF CPI, mean, change.	Mid current quarter	Y

Note: Seasonal adjustment (S.A.) or lack thereof is indicated in the fourth column. All de-meaned variables were de-meaned using a 1-year moving average. Variables obtained from the Survey of Professional Forecasters (SPF) were obtained from column 3 of their spreadsheet.

The survey also asks whether banks changed their credit spread in the past quarter, so we collected data with attention to the release data. We assume banks make decisions at the middle of each quarter and are able to access public data released near that time. They do not, however, have access to the data released near the end of the current quarter. Because banks have an asymmetric response to economic data as shown in Fig.2, exogenous variables are de-meaned by their one-year exponential moving average (EMA), with positive and negative values considered separately.

To select relevant variables from this list we used the Boruta method; an algorithm that wraps the Random Forest classification algorithm (Breiman, 2001) and "iteratively removes the features which are proved by a statistical test to be less relevant than random probes" (Kursa and Rudnicki, 2010). The random-forest approach is an ensemble method that aggregates multiple decision trees. For a classification problem, a decision tree tries to split the data into groups based on input variables (or 'features'). For a regression problem, it takes the average of all points within a group.

Random forests combine independent decision trees and randomly choose the initial variable, decision boundaries, and variables used in each node when building trees. It makes decisions by the majority vote from the end nodes of each tree. Breiman (2001) demonstrated that the results generated with this method are unbiased and have low variance. In addition, the result always converges and its accuracy is usually immune from irrelevant variables. The robustness of the random forest approach was shown by Caruana and Niculescu-Mizil (2006) who compared multiple supervised-learning algorithms including support vector machines, neural nets, random forests, and logistic regression. While each method had an edge in one or two large datasets, random forests and neural nets were found to have the best overall performance across 11 large datasets. Kane et al. (2014) used the random forest method to predict outbreaks of H5N1 and found that it out-performed ARIMA. Khaidem et al. (2016) used random forests to predict trends in stock market and obtain 85% to 95% accuracy for one-month to three-month predictions.

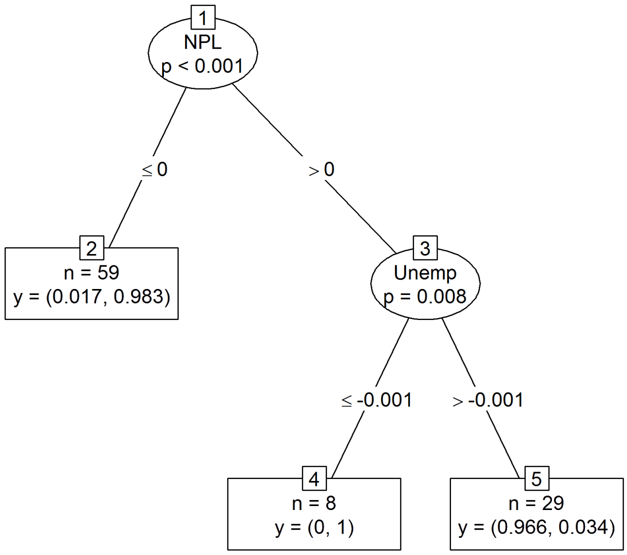

Figure 4. An Example of Decision Tree

An example of decision tree built using part of the dataset from Table 3 with one node and depth of two is shown in Fig.4. Positive values in diffusion index of credit spread are labeled as 1 (to lend) and negative values are labeled as 0 (not to lend). The decision tree starts with the change of the non-performing loan rate as an initial split; the unemployment rate is then used to further partition the result.

The result of our random-forest variable selection is shown in Table 4. In this selection process we repeat the Boruta method 20 times with random sampling and count the number of times each variable is selected either as important or weak. The variables with "positive" or "negative" modifiers are those with values above or below their EMAs and provide different signals. Credit spread (DICS), loan demand (DILD), non-performing loans below mean (NPL negative), and excessive bond premium (EBP) are constantly selected as important. Unemployment rate below mean (Unemp negative) also

Table 4. Boruta selection of simulation variables for the United States.

Variable	*Important*	*Weak*	*Variable*	*Important*	*Weak*
DICS	20	0	SPF unemp positive	3	1
Unemp positive	20	0	SPF Corp negative	2	0
NPL negative	20	0	corp negative	1	2
EBP	19	0	SPF unemp negative	1	0
NPL positive	16	2	SPF Corp positive	0	0
DILD	16	1	Unemp negative	0	0
VIX	13	5	Corp positive	0	0
CPI	12	4	SPF RGDP positive	0	0
SPF CPI	11	4	RGDP negative	0	1
NASDAQ	8	1	RGDP positive	0	0
			SPF RGDP negative	0	1

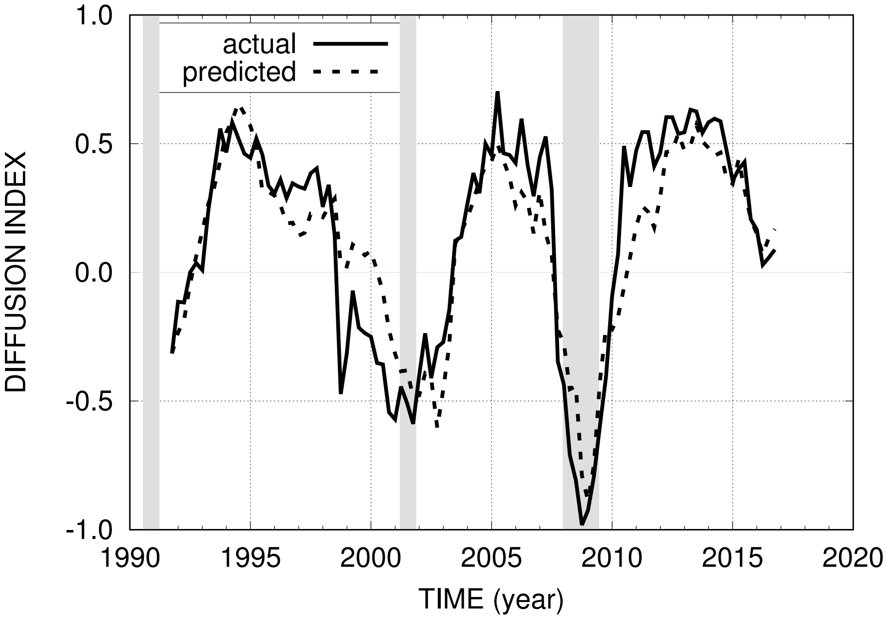

Figure 5. The observed credit spread diffusion index for the United States together with the opinion-formation model using the variables selected by the Boruta method.

ranks highly. Surprisingly, real GDP growth is not important; possibly because of its lags. We used the variables in the left column to calibrate the opinion-formation model because they have much higher importance scores.

The model coefficients of this calibration of Eq.7 are shown in the upper portion of Table 5. The diffusion index of credit spread (DICS) has a very significant effect. The excessive bond premium (EBP) and unemployment above its one year EMA also has a significant impact on lending decisions.

Our simulation of the opinion-formation model with the variables suggested by the random forest procedure is shown in Fig.5. The first 5 observations are omitted because the DI loan demand is unavailable during that time. The simulation shares the same initial value with the actual data, but subsequent values in the dynamics driver U are updated with calculated values. It captures the general trend of lending over the last two decades, but it fails at two significant moments near 1998 and 2007 for reasons that might lie in a spike in uncertainty seen in Fig.2.

The first moment was the drop of the DI credit spread in Q3 1998 from 0.143 to -0.473. This is when Russia had its financial crisis, or "Russian Flu" (Kindleberger and O'Keefe, 2011). At this time Long-Term Capital Management (LTCM) was long Russian bonds and had business ties with many major banks; it went under as the Russian market crashed and banks waited for the Fed to clear market stress. Two quarters later the credit spread rebounded to -0.071. During this period, loan demand increased and unemployment rate continued to be lower its EMA and non-performing loan rates increased above its average, suggesting the Russian Flu was responsible for this drop in the DICS.

The second moment was the significant drop in the DICS in Q3 of 2007 from 0.321 to -0.346. This was a quarter after the bank run on Countrywide Financial; Northern Rock also had bank run on September 14th 2007 (Kindleberger and O'Keefe, 2011). It appears that banks might feel financial storms well before they are headlines.

Table 5. Estimated coefficients for the opinion-formation model in the United States given by Eq.(7) with x_t corresponding to the DICS and the set of Z_i as indicated.

Parameter	*θ Estimate*	*Standard Error*	*z*	*P > |z|*	*95% C.I.*
Z_i **= Boruta-selected variables:**					
ν	0.8859	0.135	6.561	0.000	0.621 1.151
constant	0.2298	0.077	2.989	0.003	0.079 0.380
DICS	0.6213	0.137	4.534	0.000	0.353 0.890
NPL positive	-45.9799	27.395	-1.678	0.093	-99.674 7.714
NPL negative	-71.6126	24.658	-2.904	0.004	-119.942 -23.283
CPI	-26.0254	8.433	-3.086	0.002	-42.554 -9.497
EBP	-0.4983	0.141	-3.523	0.000	-0.776 -0.221
DILD	0.1102	0.168	0.656	0.512	-0.219 0.439
VIX	-1.6297	0.844	-1.932	0.053	-3.283 0.024
Unemp positive	-9.7948	2.593	-3.777	0.000	-14.877 -4.712
SPF CPI	-0.0035	0.152	-0.023	0.982	-0.302 0.295
NASDAQ	-0.2662	0.355	-0.749	0.454	-0.962 0.430
$\mathcal{LL}$	76.41				
AIC	-130.8				
BIC	-102.2				
Z_i **= all variables:**					
ν	0.5944	0.099	6.031	0.000	0.401 0.788
constant	0.0291	0.101	0.288	0.773	-0.169 0.227
DICS	0.6263	0.168	3.718	0.000	0.296 0.956
DILD	0.3454	0.232	1.490	0.136	-0.109 0.800
EBP	-0.7833	0.195	-4.008	0.000	-1.166 -0.400
VIX	-2.0566	1.075	-1.912	0.056	-4.164 0.051
NASDAQ	-0.1577	0.453	-0.348	0.728	-1.045 0.730
CPI	-26.2813	10.057	-2.613	0.009	-45.992 -6.571
SPF CPI	0.1437	0.195	0.735	0.462	-0.239 0.527
RGDP positive	48.7490	21.948	2.221	0.026	5.732 91.766
RGDP negative	-34.3133	21.856	-1.570	0.116	-77.151 8.524
SPF RGDP positive	-2.3051	24.012	-0.096	0.924	-49.367 44.757
SPF RGDP negative	-31.1695	28.640	-1.088	0.276	-87.303 24.964
Unemp positive	-10.3816	3.741	-2.775	0.006	-17.714 -3.049
Unemp negative	-7.3362	4.259	-1.722	0.085	-15.684 1.012
SPF unemp positive	3.1802	4.008	0.793	0.427	-4.675 11.036
SPF unemp negative	-3.5309	3.277	-1.077	0.281	-9.954 2.892
Corp positive	-1.2457	1.061	-1.174	0.240	-3.325 0.834
Corp negative	4.0447	1.484	2.725	0.006	1.136 6.954
SPF Corp positive	3.7323	2.456	1.519	0.129	-1.082 8.547
SPF Corp negative	1.4935	3.165	0.472	0.637	-4.709 7.696
NPL positive	-48.8259	42.971	-1.136	0.256	-133.048 35.396
NPL negative	-73.9522	31.240	-2.367	0.018	-135.182 -12.723
$\mathcal{LL}$	93.97				
AIC	-143.00				
BIC	-98.63				

Note: The model was estimated using 107 observations from 70 banks. $\mathcal{LL}$, AIC, and BIC are log-likelihood, Akaike information criteria, and Bayesian information criteria, respectively.

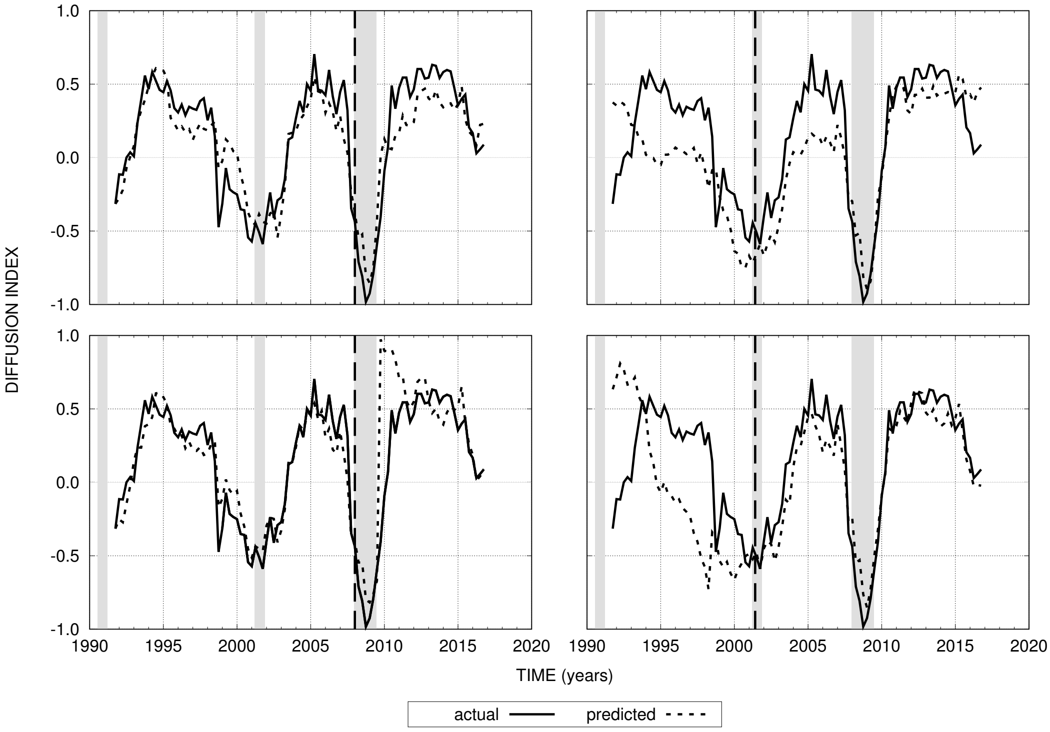

Figure 6. A comparison of simulation results in the United States using random-forest chosen variables (upper panels) and all variables (lower panels). The dashed fiducial indicates the end of the calibration set and the beginning of the simulation.

To appreciate the effect of variable reduction provided by the random forest procedure, we compare the model with all variables. The results of the model estimation are shown in the lower portion of Table 5. The coefficients seem to have expected signs with the exception of corporate profit and the DICS and EBP remain significantly above zero. However, this model appears to be over-fit and has become sensitive noise; a situation made clear when we perform cross validation of simulations that compare the random-forest reduced variable set with the complete variable set shown in Fig.6. In this analysis the forward 'prediction' used the first 70 points to estimate the model and then 'predicted' the actual diffusion index by with actual data (except the DI credit spread) for the exogenous variables. For the backward propagation, we use the last 70 points to build the model and then predicted the data (except the DI credit spread) 'backward'. The exact date of the switch was chosen to be near the financial crisis.

The reduced-variable model has relatively stable results for the forward 'prediction' shown in the upper left-hand panel of Fig.6 as the economy moves out of recession, and while the backward propagation shown in the upper left-right panel of Fig.6 is less accurate it still loosely follows the actual path. The unexpected 'Russian Flu' may have large impact here. When all variables are included the simulation becomes unstable. The forward prediction shown in the lower left-hand panel of Fig.6 fits the data before 2008 well, but explodes near 2009. The backward propagation lower right-hand panel of Fig.6 fits the data well to 2004, but the propagation to earlier times loses almost all predictive power.

3.2 Credit Dynamics in the Euro Area

The European Central Bank quarterly bank lending survey was launched in 2003 with around 90 banks in all European areas and expanded in coverage to around 140 banks in 2016. Most are large banks, but specialized small banks are also included. The sample size for each country depends on its share of loan to private non-financial sector. The survey results are aggregated at the country level and compiled to Euro Area with weights on each country's loan share (Köhler-Ulbrich et al., 2016). The survey questions are similar to those of the U.S. SLOOS, covering lending to firms and households and factors changing lending policies. They are available in the ECB's Statistical Data Warehouse. We focus on the credit spread for average loans to firms:

> *3. Over the past three months, how have your bank's terms and conditions for new loans or credit lines to enterprises changed? Please rate the overall terms and conditions for this loan category and each factor using the following scale: 1. tightened considerably, 2. tightened somewhat, 3. remained basically unchanged, 4. eased somewhat, 5. eased considerably*
>
> *(c) Your banks' loan margin (i.e. the spread over a relevant market reference rate) on average loans (wider spread = tightened, narrower spread = eased)*

We chose the data series on average loans to all firms because it is the longest. Its diffusion index (DICS), constructed by the net percent of banks that eased minus banks that tightened credit spread weighted by each country's share of outstanding loans, is shown in Fig.7; the grey bars being Euro Area recessions indicated by OECD Composite Leading Indicators. The Euro Area diffusion index behaves similarly to that of the U.S.: banks tighten lending immediately before recessions and ease credit quickly once the economy moves out of recession. An example of this is seen in the period from

2003 to 2007 when most European banks switched from tightening to easing, and in 2008 when the DICS spread dropped significantly.

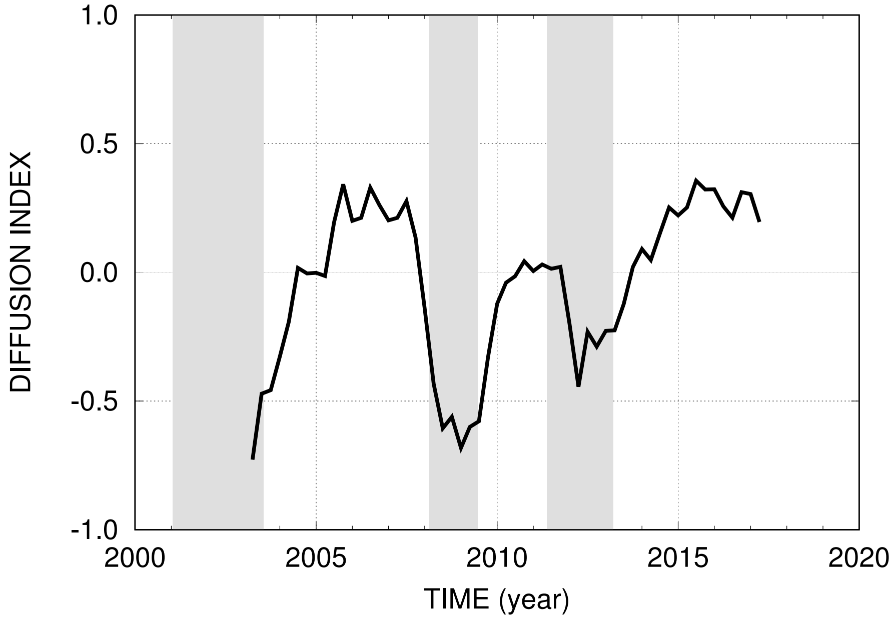

Figure 7. Euro Area Credit Spread Diffusion Index (DICS)

The ECB survey also asks banks to rank factors that affect lending polices:

4. Over the past three months, how have the following factors affected your bank's credit terms and conditions as applied to new loans or credit lines to enterprises (as defined in the notes to question 3)? Please rate the contribution of the following factors to the tightening or easing of credit terms and conditions using the following scale: 1. contributed to tighten considerably, 2. contributed to tighten somewhat, 3. contributed basically unchanged, 4. contributed to eased somewhat, 5. contributed to eased

(a) *Cost related to your bank's capital position*
(b) *Your bank's ability to access market financing*
(c) *Your bank's liquidity position*
(d) *Competition from other banks*
(e) *Competition from non-banks*
(f) *Competition from market financing*
(g) *General economic situation and outlook*
(h) *Industry or firm-specific outlook/borrowers' creditworthiness*
(i) *Risk related to the collateral demanded*

These potential drivers of lending sentiment in the Euro Area are shown together with the DICS in Fig.8. We see the same asymmetric effect on lending decisions as seen in US. When banks ease credit,

Table 6. Estimated coefficients of the minima opinion-formation models for the Euro Area.

Parameter	*Estimate*	*Standard Error*	*z*	*P > |z|*	*95% C.I.*
Model 1: Eq. (14)					
ν	2.0156	0.371	5.439	0.000	1.289 2.742
α_0	0.0053	0.017	0.317	0.751	-0.027 0.038
α_1	0.9512	0.062	15.235	0.000	0.829 1.074
$\mathcal{LL}$	40.61				
AIC	-77.22				
BIC	-73.17				
Model 2: Eq. (15)					
ν	2.0352	0.378	5.382	0.000	1.294 2.776
α_0	0.0176	0.029	0.607	0.544	-0.039 0.074
α_1	0.8795	0.152	5.769	0.000	0.581 1.178
α_2	0.9912	0.098	10.125	0.000	0.799 1.183
$\mathcal{LL}$	40.75				
AIC	-75.49				
BIC	-69.41				

Note: The model was estimated using 56 observations from 140 banks. $\mathcal{LL}$, AIC, and BIC are log-likelihood, Akaike information criteria, and Bayesian information criteria, respectively.

competition – especially from other banks – is the main force as seen in the upper-left panel of Fig.8. This explains the sharp increase in the DICS near 2005. The other reported factors shown in Fig.8 have almost no impact on the upside. When banks tighten lending the general economic outlook tracks the DICS well as shown in the upper-right panel. Industrial problems and borrowers' creditworthiness seen in the lower panels of this Figure are also important in tightening credit but have almost no impact in easing. Surprisingly, the liquidity and capital positions, which are generally regarded as important in lending decisions, have far less importance as indicated in the middle panels of Fig.8.

We applied the opinion-formation model to the Euro Area's diffusion index of credit spread, beginning with Models 1 and 2 from the U.S. analysis to examine potential asymmetry in the peer effect. The results shown in Table 6 are quite similar to those of the U.S., but from an information perspective (AIC and BIC) Model 1 outperforms Model 2, suggesting that separating this diffusion index by sign may not be an optimal choice. The simulation of Model 2 is shown in Fig.9 where the DI credit spread in the dynamics driver U is updated by actual data. The simulation misses the beginning of the Great Recession in 2007-8, but otherwise tracks the actual data well.

The candidate variables for the expanded opinion-formation model of the Euro Area shown in Table 7 were selected based on our experience with the U.S. data. In the Euro Area, the VIX, 'fear index', is replaced by the EURO STOXX 50 obtained from the Factset database. Forecast data for GDP, unemployment rate, and inflation rate was sourced from the ECB's SPF website. Corporate profit in the U.S. had little impact on lending decisions so it was excluded from the Euro data. The non-performing loan rate for the Euro Area is available, but it starts in Q2 2015 and is therefore not considered here. Excessive bond premium for the Euro Area is absent because bank level information is not publicly available.

The variable importance obtained from the Boruta method is shown in Table 8. In the absence of excessive bond premium and non-performing loan rates, the unemployment rate becomes an important variable. Real GDP, with two-period lags (because of the release date) and the forecast are less important. As before we chose variables in the left column because of their relevance.

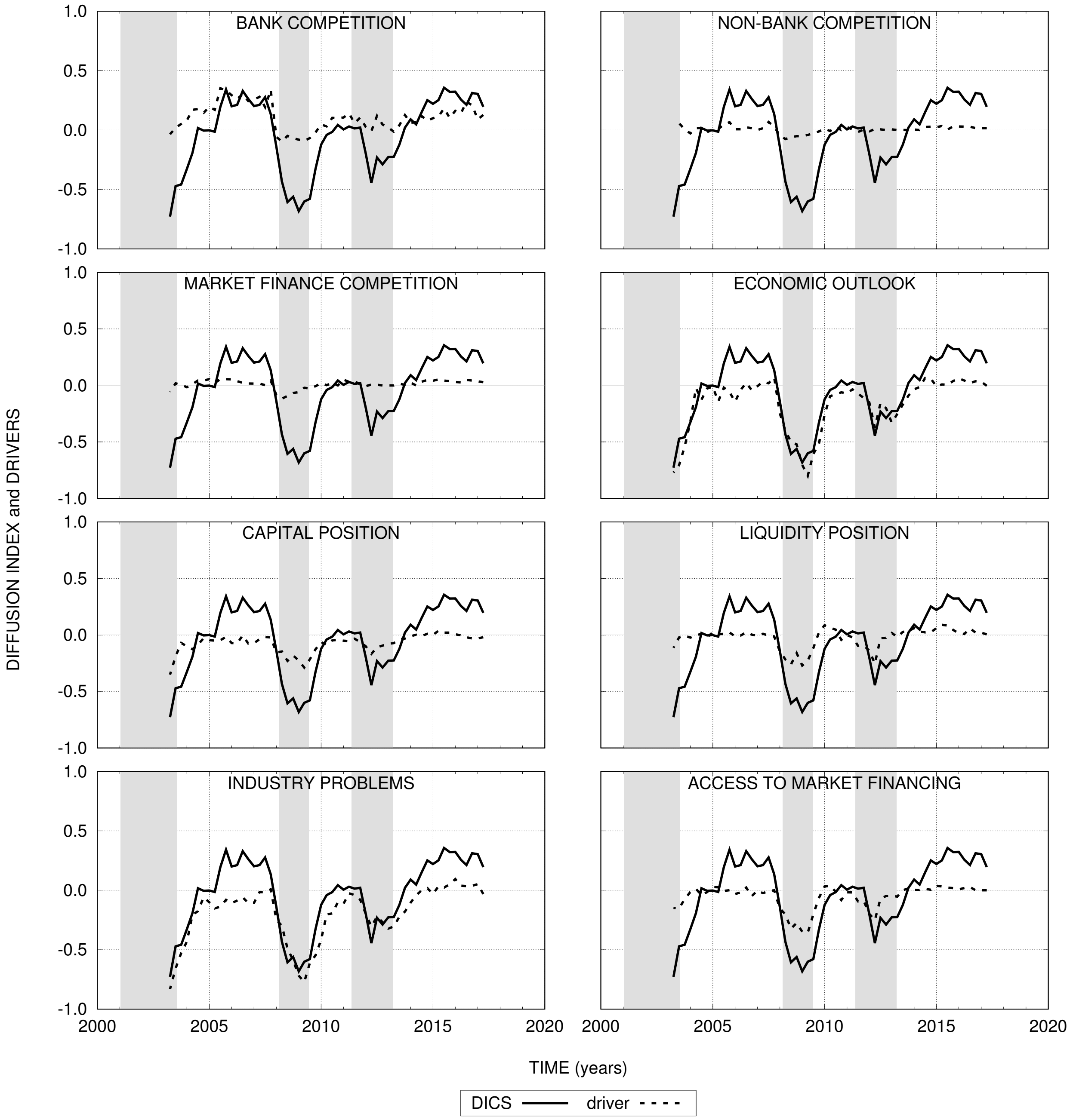

Figure 8. The credit spread diffusion index together with potential drivers of the diffusion index in the Euro Area.

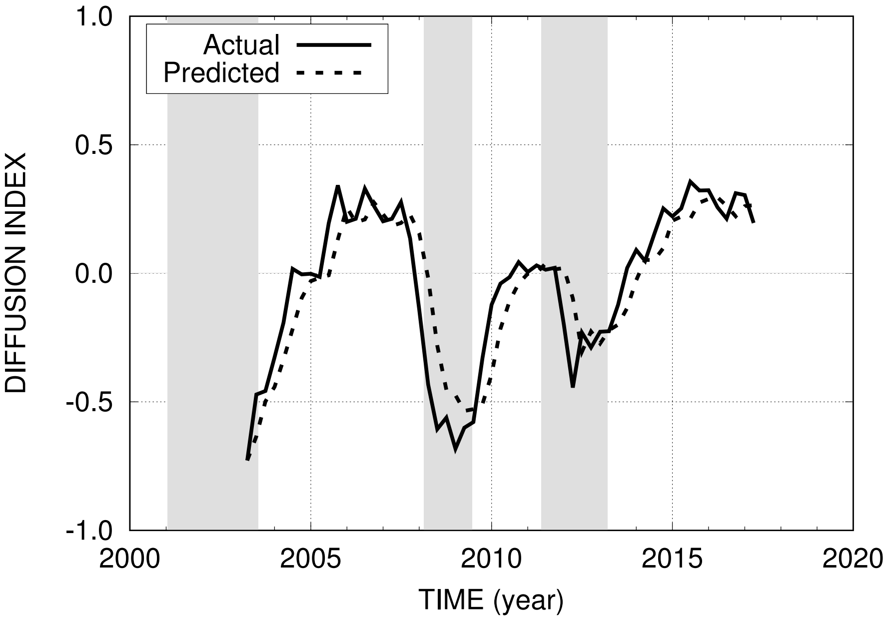

Figure 9. Simulation of the opinion-formation model for the Euro Area given by Eq. (15).

Table 7. Candidate variables for the opinion-formation model in the Euro Area given by Eq. (7).

Variable	*Description*	*Release Date*	*S.A.*
DICS	Diffusion index of credit spread, average loans.	Mid next quarter	N
DILD	Diffusion index of loan demand.	Mid next quarter	N
STOXX	Euro STOXX 50 volatility index. Quarter average, change.	Daily	N
RGDP	RGDP growth rate. De-meaned.	End of next quarter	Y
Unemp	Unemployment rate Euro Area, percent change. De-meaned.	Monthly	Y
SPF RGDP	SPF RGDP growth rate, next year. De-meaned.	Mid current quarter	Y
SPF Unemp	SPF unemp rate, percent change, mean. De-demeaned.	Mid current quarter	Y
Inflation	Overall index (HICP). Percent change.	Monthly	N
SPF Inflation	SPF CPI, next year. Percent change.	Mid current quarter	N

Note: Seasonal adjustment (S.A.) or lack thereof is indicated in the fourth column. All de-meaned variables were de-meaned using a 1-year moving average. Variables obtained from the Survey of Professional Forecasters (SPF) were obtained from column 3 of their spreadsheet.

Table 8. Boruta selection of simulation variables for the Euro Area.

Variable	*Important*	*Weak*	*Variable*	*Important*	*Weak*
DICS	20	0	Unemp negative	3	3
SPF Unemp positive	17	2	SPF Unemp negative	1	1
DILD	16	2	RGDP lag2 positive	0	0
Unemp positive	12	3	RGDP lag2 negative	0	0
STOXX	13	4	SPF RGDP positive	0	0
Inflation	6	2	SPF RGDP negative	0	0
SPF Inflation	3	8			

Table 9. Estimated coefficients for the opinion-formation model in the Euro Area given by Eq. (7) with x_t corresponding to the DICS and the set of Z_i as indicated.

Parameter	*θ Estimate*	*Standard Error*	*z*	*P > |z|*	*95% C.I.*
Z_i **= Boruta-selected variables:**					
ν	1.1854	0.225	5.259	0.000	0.744 1.627
constant	0.1879	0.194	0.970	0.332	-0.192 0.567
DICS	0.5242	0.188	2.788	0.005	0.156 0.893
DILD	0.0797	0.202	0.395	0.693	-0.316 0.475
STOXX	0.3259	0.294	1.110	0.267	-0.250 0.902
Inflation	-9.5982	5.539	-1.733	0.083	-20.454 1.258
Unemp positive	-3.6388	4.773	-0.762	0.446	-12.994 5.716
Unemp negative	-0.2153	5.103	-0.042	0.966	-10.218 9.787
SPF Unemp positive	-5.0701	2.889	-1.755	0.079	-10.733 0.592
SPF Unemp negative	-0.6924	2.122	-0.326	0.744	-4.851 3.466
SPF Infla	1.2303	15.288	0.080	0.936	-28.735 31.195
$\mathcal{LL}$	55.91				
AIC	–91.82				
BIC	-71.74				
Z_i **= all variables:**					
ν	2.2894	0.425	5.382	0.000	1.456 3.123
constant	0.0666	0.104	0.641	0.522	-0.137 0.270
DICS	0.8334	0.099	8.442	0.000	0.640 1.027
DILD	-0.0159	0.107	-0.148	0.882	-0.226 0.194
STOXX	0.1780	0.155	1.148	0.251	-0.126 0.482
Inflation	-3.7667	2.940	-1.281	0.200	-9.529 1.996
SPF Infla	-0.0845	8.002	-0.011	0.992	-15.768 15.599
RGDP lag2 positive	0.4055	525.526	0.001	0.999	-1029.606 1030.417
RGDP lag2 negative	-0.9045	2193.465	-0.000	1.000	-4300.017 4298.208
Unemp positive	-2.3672	2.493	-0.949	0.342	-7.254 2.520
Unemp negative	0.3466	2.676	0.130	0.897	-4.898 5.591
SPF RGDP positive	0.4055	525.526	0.001	0.999	-1029.606 1030.417
SPF RGDP negative	-0.9045	2193.464	-0.000	1.000	-4300.015 4298.206
SPF Unemp positive	-2.4533	1.500	-1.635	0.102	-5.394 0.487
SPF Unemp negative	-0.8982	1.126	-0.798	0.425	-3.106 1.309
$\mathcal{LL}$	40.75				
AIC	-75.49				
BIC	-69.41				

Note: The model was estimated using 56 observations from 140 banks. $\mathcal{LL}$, AIC, and BIC are log-likelihood, Akaike information criteria, and Bayesian information criteria, respectively.

The model estimation for selected variables in the upper portion of Table 9 gives consistent results and the variables have the expected signs. However, the unemployment rate may have less explanatory power because of the wide range of unemployment rates across Europe.

3.3 Credit Dynamics in Japan

As we did for the U.S. data, we also compare the full variable estimation as a stability check. The full data estimation is shown in the lower portion of Table 9, and we also compare the forward prediction and backward propagation shown in Fig.11. For the forward prediction shown in the left-hand panels of Fig.11 we use the data before Q3 2013 to estimate the parameters and then update real data (except DI credit spread) to compute the prediction. For the backward propagation shown in the right-hand panels of Fig.11 we use the data after Q3 2006 to build the model and use that date as the

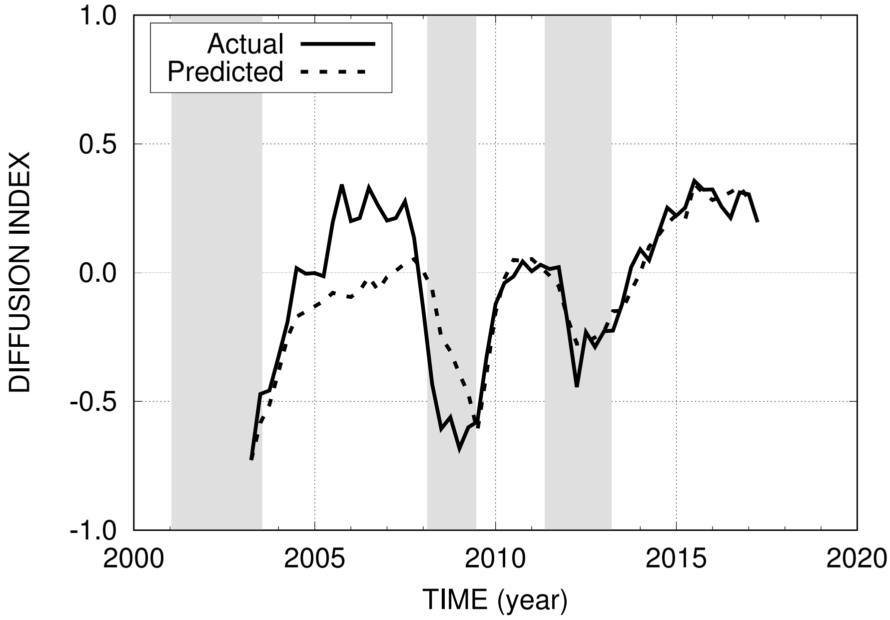

Figure 10. Simulation of the opinion-formation model for the Euro Area given by Eq. (15).

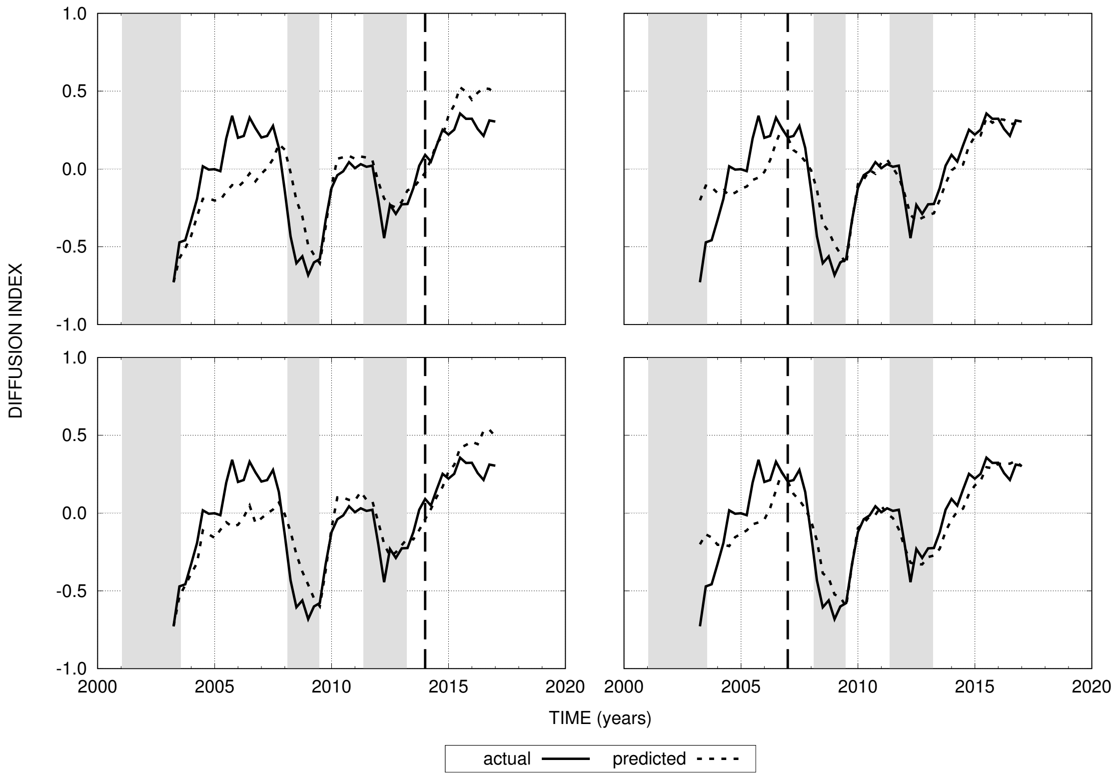

Figure 11. A comparison of simulation results in the Euro Area using random-forest chosen variables (upper panels) and all variables (lower panels). The dashed fiducial indicates the end of the calibration set and the beginning of the simulation.

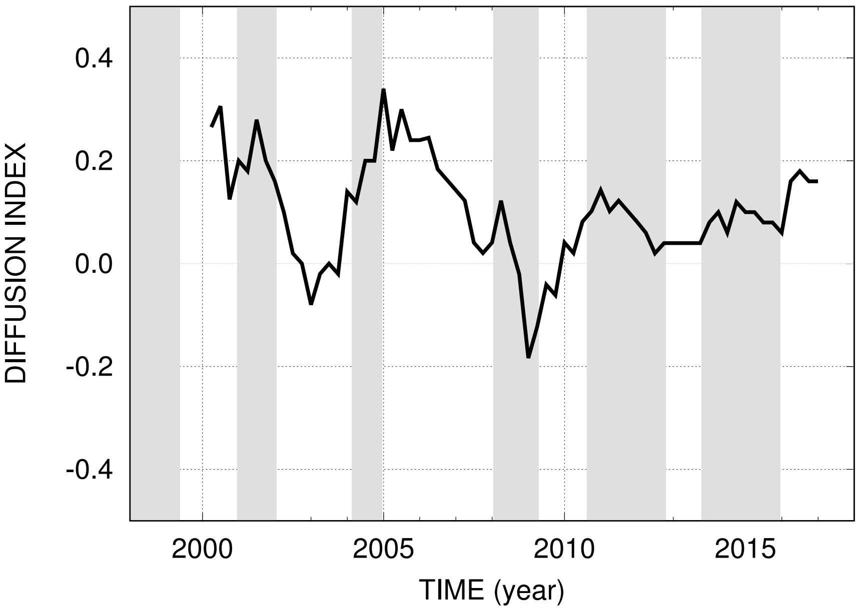

Figure 12. The credit spread diffusion index (DICS) for Japan.

initial position. Then we backward update the data to compute the prediction. The forward prediction is relatively stable but the back propagation is very unstable. A potential explanation is the data length: the U.S. requires around 60 points to obtain stable parameters, but in these two simulations we only used 40 points for parameter estimation. The Euro Area data may be too short at this time to give a stable estimation for the extended model.

The structure and questions of the Bank of Japan (BOJ) lending survey are very similar to that of the United States. Credit spread data, however, is not available so we extracted the needed data from BOJ reports and calculated the DICS by taking the net percent difference of banks that eased less those that tightened. The resulting DICS for high-rating firms is shown in Fig.12; the gray bars are OECD based recession period for Japan. The DICS of Japan differs significantly from those of the U.S. and the Euro Area in terms of its dynamic range. The DICS of Japan has a materially smaller dynamic range; particularly on the downside where its magnitude is less than 20% of those in the U.S. and the Euro Area.

The reasons for changing lending policy were extracted from Japanese-language survey reports. The BOJ only ask reasons for changing credit standards, but its credit standard DI has a very similar shape to the credit spread diffusion index so we use these data to approximate reasons for changing the credit spread.

> *If your bank has eased its credit standards for loans to firms over the past three months (as described in question 7), what were the important factors that led to the change? (Please rate each possible reason using the following scale: 3 = important, 2 = somewhat important, 1 = not important.)*
>
> *(a) An improvement (deterioration) in your bank's asset portfolio.*
>
> *(b) A more (less) favorable or less (more) uncertain economic outlook.*

Table 10. Estimated coefficients for the opinion-formation model for Japan given by Eq. (14).

Parameter	*Estimate*	*Standard Error*	*z*	*P > |z|*	*95% C.I.*
ν	0.1646	0.032	5.156	0.000	0.102 0.227
α_0	0.2424	0.128	1.888	0.059	-0.009 0.494
α_1	-1.6313	0.989	-1.649	0.099	-3.570 0.307
$\mathcal{LL}$	94.71				
AIC	-185.4				
BIC	-181.0				

Note: The model was estimated using 67 observations from 50 banks. $\mathcal{LL}$, AIC, and BIC are log-likelihood, Akaike information criteria, and Bayesian information criteria, respectively.

(c) An improvement (worsening) in industry or firm specific problems.

(d) Competition from other banks.

(e) Competition from non-banks.

(f) Competition from capital markets.

(g) Risk tolerance.

These potential drivers of lending sentiment in Japan are shown together with the DICS in Fig.13. The results are similar to those in the U.S. and the Euro Area. Competition from other banks has the strongest correlation with banks easing credit. Competition does not, however, appear to be a driver of credit tightening. For example, during the 2008 recession when most banks tightened credit the competition terms – whether from other banks, non-banks, or the capital markets – simply went to zero, not negative. Uncertainty has the second highest correlation with the DICS when credit is being eased. Lack of it drives lending during the boom from 2004 to 2007. During the financial crisis associated with the Great Recession in 2008 negative uncertainty also had a very strong effect. The asset portfolio of banks also had a negative impact during the Great Recession as loan repayment became less certain. The state of industry problems drives easing of credit but seems to have little impact on the downside. This may be due to a bank looking to others to ease while having ample evidence in the own non-performing loan book when times get tough to justify tightening.

Applying the opinion-formation model to Japan's DICS reveals that peer effects have a much less impact than in the U.S. and the Euro Area. In particular, as shown in Table 10, the coefficient for the DICS is negative; opposite that of the U.S. and Euro Area. The corresponding simulation shown in Fig.14 does not track the observed DICS well. The simulation begins two-years later because business ad employment forecasts began in 2004.

Economic variables – selected based upon our experience with the U.S. and Euro Area – are shown in Table 11. The TANKAN, a short-term economic survey of enterprises beginning Q2 2004, was used as a proxy for banks' expectations. Part of the survey asks banks for their opinions on the future business and employment conditions. Their diffusion index (favorable minus unfavorable) is listed in the BOJ database. The VIX equivalent for Japan is the VXJ which follows the CBOE method for calculating the VIX. Non performing and bankruptcy loans are from Japan Financial Service Agency (FSA) which monitors bad loans among banks. It classifies three types of bad loans : (i) "bankrupt or de facto bankrupt" includes loans from borrowers who are legally bankrupt; (ii) "doubtful loans" are from borrowers who are known in financial difficulties and are not expected to pay back principles

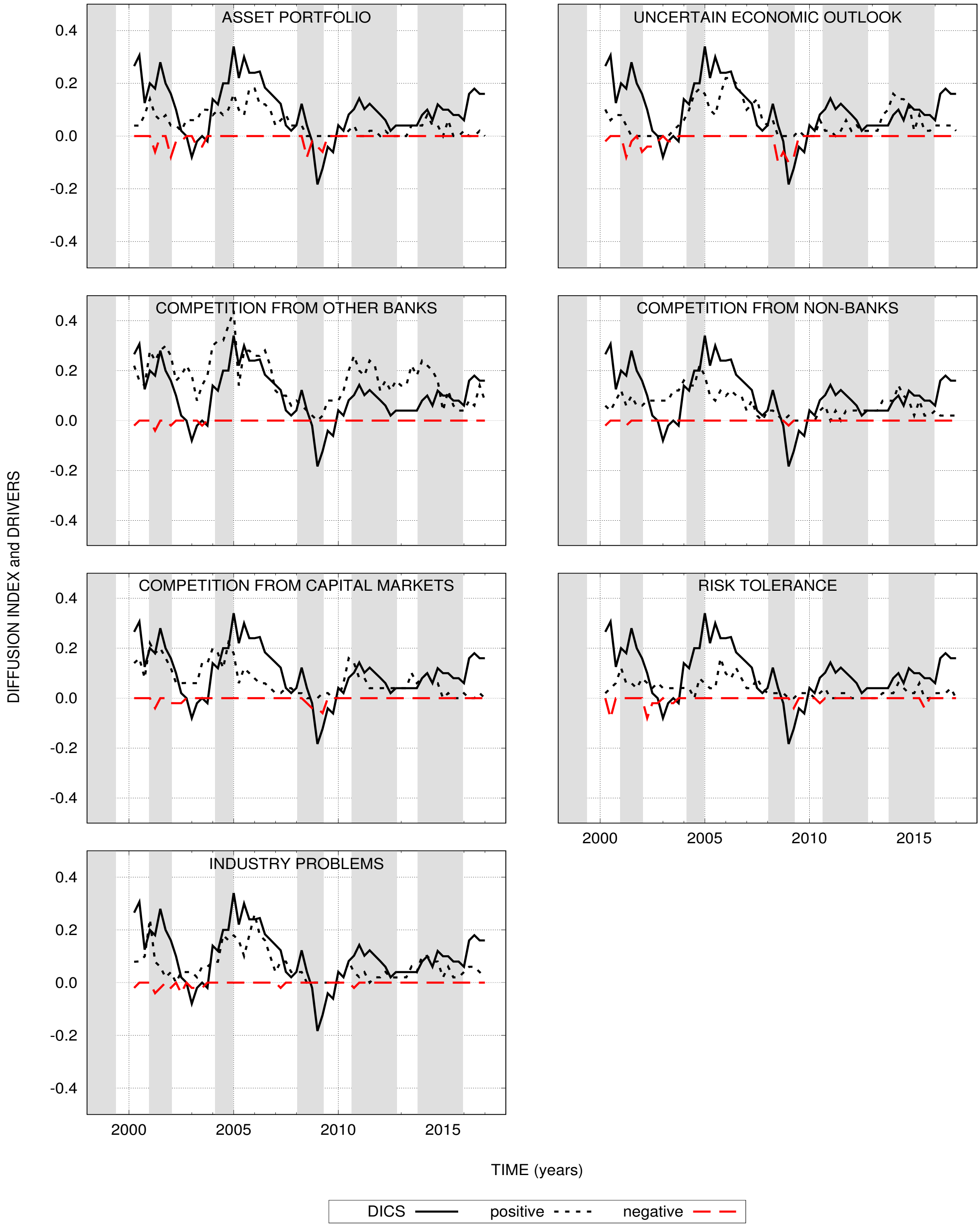

Figure 13. The credit spread diffusion index together with potential drivers of the diffusion index in Japan.

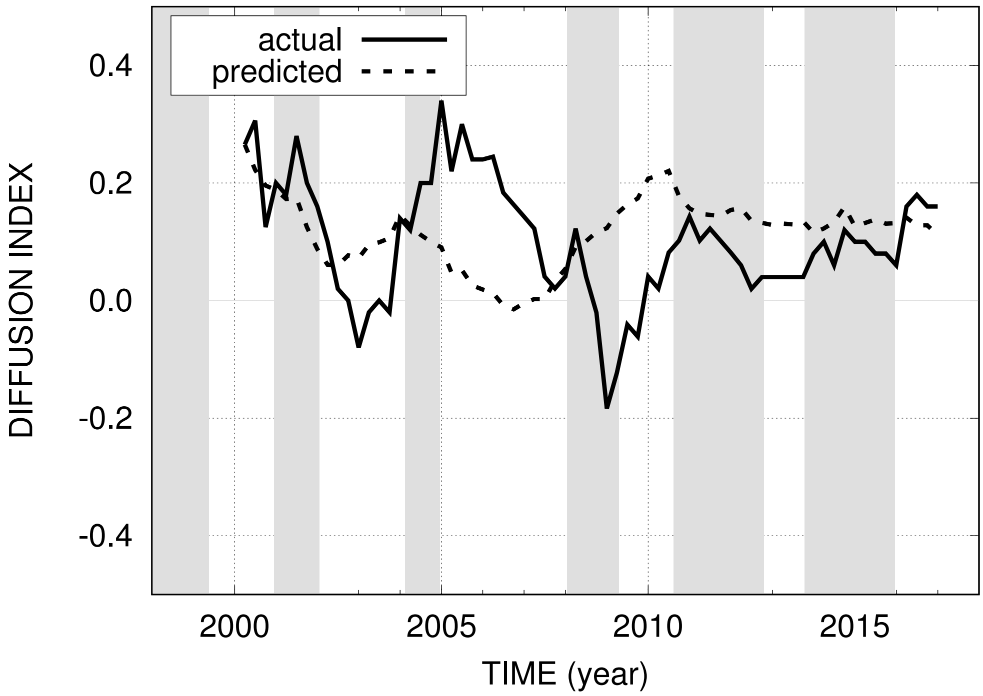

Figure 14. The observed credit spread diffusion index for Japan together with the simulation result using opinion-formation model given by Eq. (14).

Table 11. Candidate variables for the opinion-formation model of Japan given by Eq. (7).

Variable	*Description*	*Release Date*	*S.A.*
DICS	Diffusion index of credit spread for high rating firms.	Mid next quarter	N
DILHRD	Diffusion index of loan demand from high-rated firms.	Mid next quarter	N
Business Forecast	TANKAN Business Forecast from Banks.	Mid next quarter	Y
Employment Forecast	TANKAN Employment Forecast from Banks.	Mid next quarter	N
NPL	Non-performing loans. De-meaned.	Quarterly	N
Bankruptcy	Loans of debtors in bankruptcy. De-meaned.	Quarterly	N
VXJ	The VIX of Japan. Quarter average, change.	Daily	N
Nikkei 225	Nikkei Stock Average. Percent change.	Daily	N
Inflation	All item. Percent change.	Monthly	N
Unemp	Unemployment rate. Percent change. De-meaned.	Monthly	Y
RGDP	Real GDP, 2 lags. Percent change.	Mid next quarter	Y

Note: Seasonal adjustment (S.A.) or lack thereof is indicated in the fourth column. All de-meaned variables were de-meaned using a 1-year moving average.

Table 12. Boruta selection of simulation variables for Japan.

Variable	*Important*	*Weak*	*Variable*	*Important*	*Weak*
DICS	20	0	Inflation	1	3
Bankrupt negative	20	0	NPL positive	0	0
NPL negative	19	1	Employment Forecast	0	0
Bankrupt positive	19	0	RGDP lag2 positive	0	0
VJX	11	6	RGDP lag2 negative	0	4
Business Forecast	10	3	Unemp positive	0	0
Nikkei	6	3	Unemp negative	0	0
DIHRLD	3	2			

and interest on time; (iii) "special attention" loans are past due by more than 90 days (Kawashima and Nakabayashi, 2014). The data are reported by banks' self-assessment, but the FSA also publishes its inspection on major banks' bad loans. We group the non-performing loans by type (ii) and (iii) and choose bankruptcy by type (i) from major banks. The dataset is annual before 2004 and semi-annual after 2004, so we impute their missing quarterly data by cubic spline.

Our analysis of these variables using the Boruta method is shown in Table 12. Bankruptcy loans, both positive and negative, rank high in importance. Non-performing loans below their 1-year EMA (negative) also rank highly but the positive side does not, which raises the interesting question: are banks confident that their large-firm borrower will pay back loans even though they are in arrears? Banks' consensus on next quarter business matters, but their estimation on employment doesn't. Real GDP and unemployment are also not classified as important. We use variables on the left column of Table 12 because of their high importance scores.

The model estimation for these selected variables is given in the upper portion of Table 13 and shows that the coefficient for real GDP and for the VJX are both significant and have the expected sign. The inclusion of these extra variables, however, did not resolve the issue of the unexpected negative sign of the DICS.

The simulation using the variables selected by the random forest and Boruta method shown in Fig.15 shows remarkable improvement over the simpler model shown in Fig.14 and now tracks the observed DICS with the fidelity seen in the U.S. and the Euro Area. Extending the opinion-formation model for Japan to all variables as shown in the lower portion of Table 13 and we compare the predictive power of the multi-variable versions of the opinion-formation model for Japan in Fig.16. For forward prediction shown in the left-hand column of panels, we use data before Q2 2014 to estimate model and update the real data (except DI credit spread). For the backward propagation show in the right-hand column of panels, of Fig.16 we use Q4 2006 as the initial value and data afterward to estimate the model, then we backward update the real data (except DI credit spread); DI credit spread being updated by the calculated value. Including more variables via the Boruta method make the prediction better as shown in the upper panels of Fig.16. Including all variables, however, does not as shown in the lower panels of Fig.16.

Finally, to emphasize the importance of bad loans in lending decisions, we show the results of the extended opinion-formation model calibrated with and without bad loans in Fig.17. The upper row of panels show the results for the Boruta-selected variables (left) and all variables (right) when bad loans are included in the calibration as discussed above. The lower row of panels show the same when bad loans are excluded from the calibration. When bad loans are included in the calibration the

Table 13. Estimated coefficients for the opinion-formation model in Japan given by Eq. (7) with x_t as the DICS and the set of Z_i as indicated.

Parameter	*θ Estimate*	*Standard Error*	*z*	*P > |z|*	*95% C.I.*
Z_i = Boruta-selected variables:					
ν	0.0617	0.017	3.714	0.000	0.029 0.094
constant	1.0870	0.369	2.943	0.003	0.363 1.811
DICS	-11.5044	3.047	-3.776	0.000	-17.476 -5.532
VJX	-2.1154	1.512	-1.399	0.162	-5.079 0.849
Business Forecast	-0.7744	1.137	-0.681	0.496	-3.002 1.454
NPL negative	-281.0070	99.885	-2.813	0.005	-476.779 -85.235
Bankrupt positive	-4770.6882	1234.858	-3.863	0.000	-7190.964 -2350.412
Bankrupt negative	-1024.3461	1032.815	-0.992	0.321	-3048.627 999.934
DIHRLD	1.0874	0.810	1.342	0.180	-0.501 2.676
Nikkei	-2.8763	1.782	-1.614	0.107	-6.369 0.617
$\mathcal{LL}$	92.10				
AIC	-166.2				
BIC	-149.0				
Z_i = all variables:					
ν	0.0394	0.013	3.120	0.002	0.015 0.064
constant	2.4020	0.760	3.159	0.002	0.912 3.892
DICS	-18.1807	4.302	-4.226	0.000	-26.613 -9.749
DIHRLD	1.6611	1.240	1.339	0.180	-0.770 4.092
NPL positive	-1884.1503	1226.587	-1.536	0.125	-4288.217 519.916
NPL negative	-416.8001	131.756	-3.163	0.002	-675.037 -158.563
Bankrupt positive	-5296.8567	1620.667	-3.268	0.001	-8473.305 -2120.408
Bankrupt negative	-2717.2816	1327.687	-2.047	0.041	-5319.500 -115.063
Business Forecast	-4.5936	3.222	-1.426	0.154	-10.909 1.722
Employment Forecast	-4.6651	4.180	-1.116	0.264	-12.857 3.527
Inflation	-1.6268	6.379	-0.255	0.799	-14.129 10.875
VJX	-4.9289	2.006	-2.457	0.014	-8.860 -0.997
Nikkei	-5.3658	2.236	-2.400	0.016	-9.748 -0.983
RGDP lag2 positive	24.8723	44.804	0.555	0.579	-62.942 112.687
RGDP lag2 negative	82.4556	63.023	1.308	0.191	-41.067 205.978
Unemp positive	-30.9723	16.668	-1.858	0.063	-63.641 1.696
Unemp negative	0.4540	10.882	0.042	0.967	-20.875 21.783
$\mathcal{LL}$	96.94				
AIC	-161.9				
BIC	-131.3				

Note: The model was estimated using 51 observations from 50 banks. $\mathcal{LL}$, AIC, and BIC are log-likelihood, Akaike information criteria, and Bayesian information criteria, respectively.

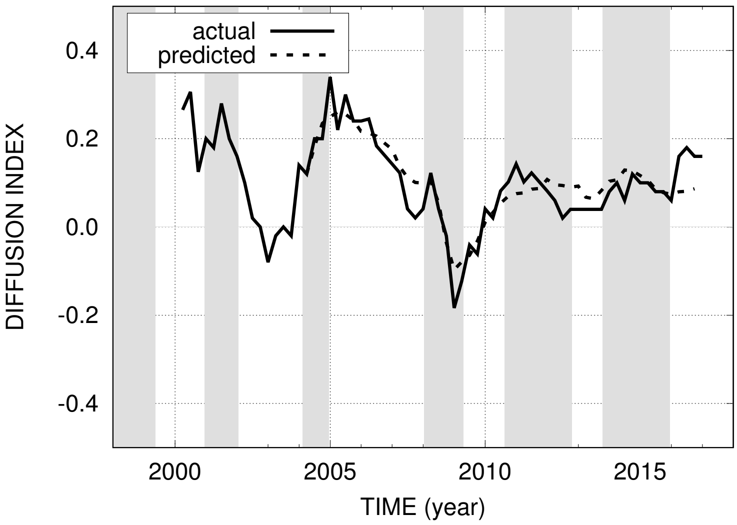

Figure 15. The observed credit spread diffusion index for Japan together with the simulation result using opinion-formation model with the variables selected by the Boruta method.

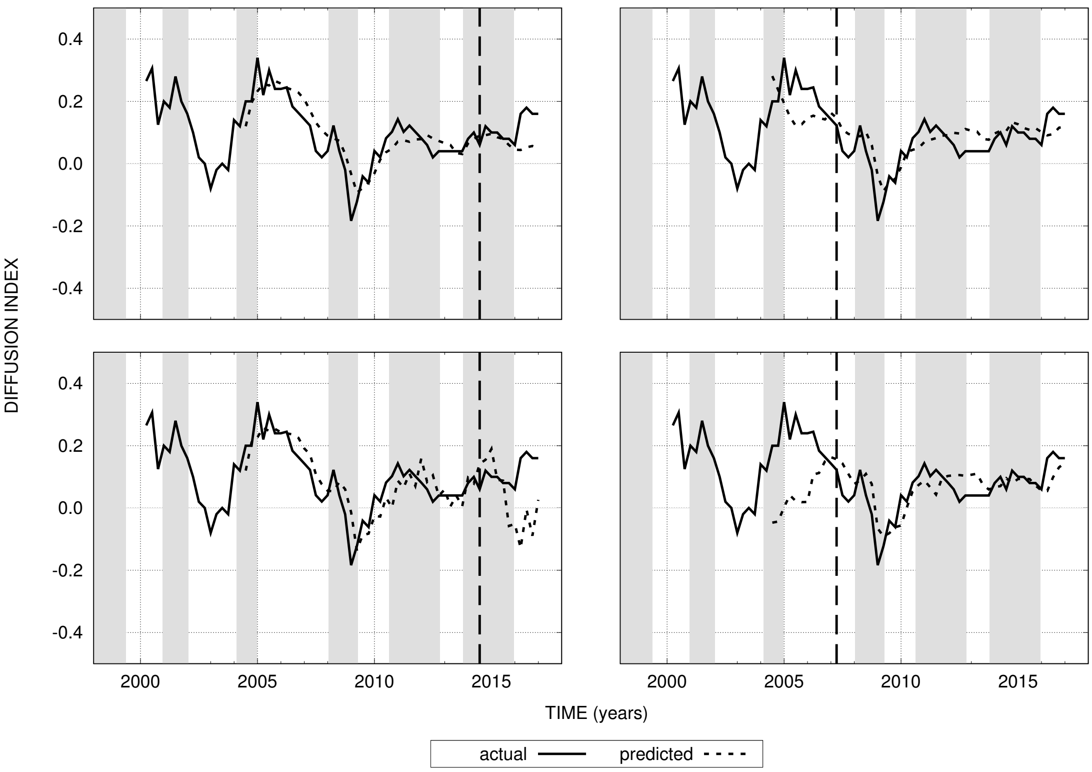

Figure 16. A comparison of simulation results in Japan using random-forest chosen variables (upper panels) and all variables (lower panels). The dashed fiducial indicates the end of the calibration set and the beginning of the simulation.

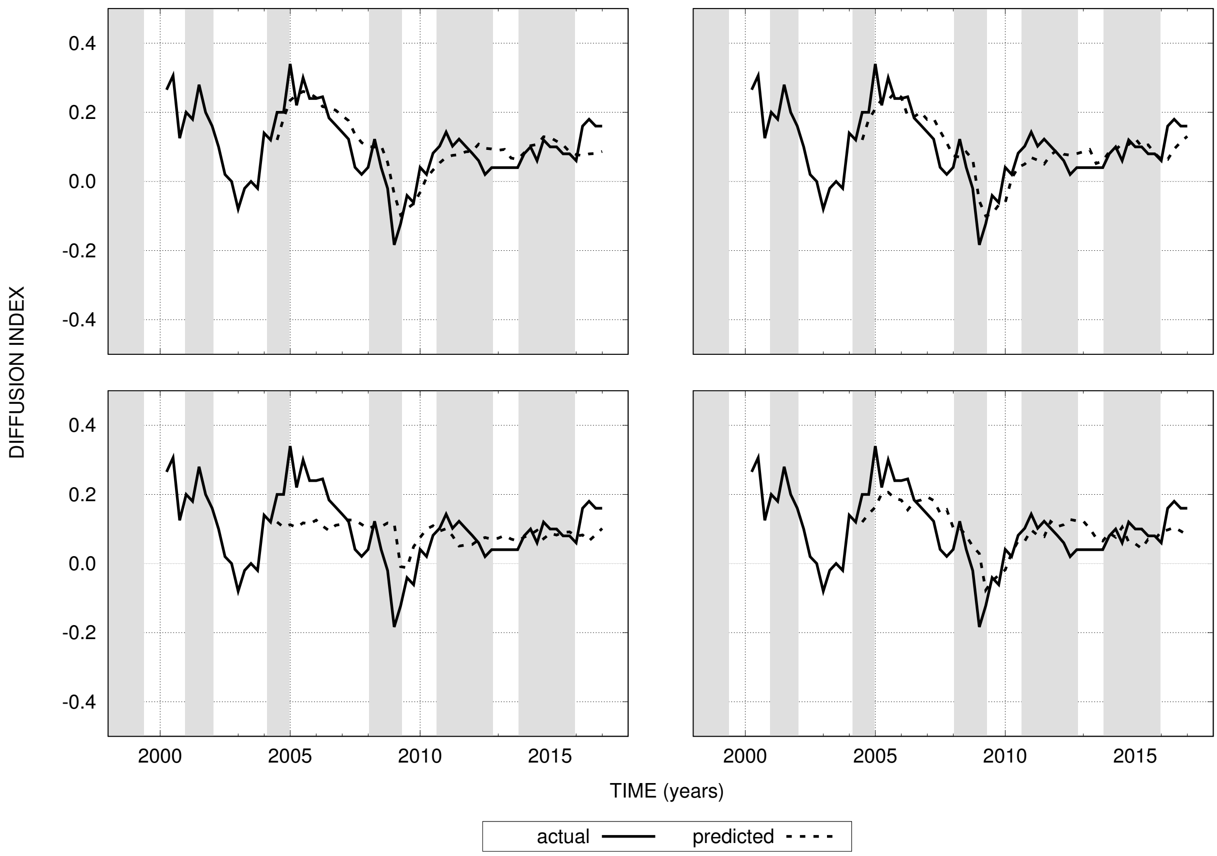

Figure 17. A comparison of simulation results in Japan using random-forest chosen variables (left panels) and all variables (right panels) when bad loans are included (upper panels) and when bad loans are excluded (lower panels).

opinion-formation model performs well; whether using the Boruta-chosen variables or all the variables. When bad loans are not included the performance of the model degrades significantly, with the Great Recession underestimated and the credit-easing cycle of 2004-2007 either underestimated or missed completely.

4. Discussion and Summary

In this paper we applied the opinion-formation model of Ghonghadze and Lux (2016) for the credit spread diffusion index (DICS) to the Euro Area and Japan and introduced the use of machine-learning techniques for variable selection. We validated our approach by comparing to the results of Ghonghadze and Lux (2016) for the U.S. as an important first step in our analysis. We find the banks react differently to economic factors for changing lending policies. Uncertainty and competition are two main factors that act differentially: the former in bad times and the latter in good times. Capital position and loan liquidity for banks in the U.S. and the Euro Area were found to have importance in the formation of lending decisions while in Japan we find that bad loans – those from firms in bankruptcy and non-performing loans – play a key role. In all regions the state of the stock market is found to be important in lending decisions.

The credit spread diffusion index displays momentum. Banks tighten the spread one or two quarters before economic downturn or unexpected event such as 'Russian Flu', and banks lend more if in the the past quarter they did so until uncertainty dominates the market; at which point liquidity decreases substantially. The success of the opinion-formation model in the DICS simulations supports the long-held notion that banks do consider their peers' lending opinions in addition to their own rational risk calculation when deciding whether to extend credit.

Interestingly, real GDP ranks very low in terms of importance. According to Okun's law, both unemployment and real GDP should have similar explanatory power. From a Post-Keynesian point of view, however, this finding is not unexpected. Loans are underwritten in nominal terms, so the price level matters. The substantial variation in economic situation across the Euro Area suggests that the promising results we found at the aggregate level here may extend to the country level as well.

The formal introduction of social psychology into the assessment of lending-opinion formation provides psychologically sound microfoundations for the understanding of the dynamics of the financial business cycle. This is critically important because suboptimal lending decisions driven largely by the actions of other banks remains remarkable robust as noted by John Mack of Morgan Stanley (Moore, 2009)

> "I missed a piece of business. I can live with that, but as soon as I hung up the phone someone else put up 10 times leverage. We cannot control ourselves. You have to step in and control the Street."

to advances in economic risk measurement and management since Keynes' observations of banking social psychology in 1931. The opinion-formation model presented in this paper and in those we have cited shows promise as a framework for a coherent synthesis of psychology and economics as it pertains to the financial business cycle.

Acknowledgement

We thank Professor Thomas Lux and Dr. Jaba Ghonghadze for helpful discussions concerning this work in general and for sharing a dataset that allowed us to confirm validation of our reproduction of their results in particular. We also thank Noah Golmant for helpful conversations on aspects of the simulation and data handling.

Conflict of Interest

All authors declare no conflicts of interest in this paper.

References

Altavilla C, Darracq Paries M, Nicoletti G, et al. (2015) Loan supply, credit markets and the euro area financial crisis. *ECB Working Paper 1861*, European Central Bank.

Bouwman CH, Malmendier U (2015) Does a bank's history affect its risk-taking? *Am Econ Rev* 105: 321-325.

Breiman L (2001) Random forests. *Mach Learn* 45: 5-32.

Caruana R, Niculescu-Mizil A (2006) An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on Machine learning*, ACM, 161-168.

Favara G, Gilchrist S, Lewis K, Zakrajsek E (2016) Recession risk and the excess bond premium. *FEDS Notes*, Board of Governors of the Federal Reserve System.

Ghonghadze J, Lux T (2012) Modelling the dynamics of EU economic sentiment indicators: an interaction-based approach. *Appl Econ* 44: 3065-3088.

Ghonghadze J, Lux T (2016) Variation of lending standards and the financial instability hypothesis. *FinMaP Working Paper*, Kiel University.

Gilchrist S, Zakraj (2012) Credit spreads and business cycle fluctuations. *Am Econ Rev* 102: 1692-1720.

Granovetter M (2005) The impact of social structure on economic outcomes. *J Econ Perspect* 19: 33-50.

Hawkins RJ (2011) Lending sociodynamics and economic instability. *Phys A: Stat Mech its Appl* 390: 4355-4369.

Helbing D (2010) *Quantitative Sociodynamics: Stochastic Methods and Models of Social Interaction Processes*, Berlin: Springer.

Holmes A (1969) Operational constraints on the stabilization of money supply growth. *Controlling Monetary Aggregates, (Boston MA: Federal Reserve Bank of Boston)* 65-77.

Kane MJ, Price N, Scotch M, Rabinowitz P (2014) Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC Bioinf* 15: 276.

Kawashima T, Nakabayashi M (June, 2014) Non-performing loan reduction with regulatory transition: A Japanese experience, 1998–2013. *FSA Institute Discussion Paper Series DP2014-2*, Financial Research Center (FSA Institute) Financial Services Agency.

Keynes JM (1931) The consequences to the banks of the collapse in money values (Aug. 1931). In *Essays in Persuasion*, New York: W. W. Norton & Company, book published in 1963, 168-180.

Keynes JM (1937) The general theory of employment. *Q J Econ* 51: 209-223.

Khaidem L, Saha S, Dey SR (2016) Predicting the direction of stock market prices using random forest. *arXiv preprint arXiv:1605.00003*.

Kindleberger CP, O'Keefe R (2011) *Manias, Panics, and Crashes*, Springer.

Köhler-Ulbrich P, Hempell HS, Scopel S (2016) The euro area bank lending survey: Role, development and use in monetary policy preparation. *Occasional Paper 179*, European Central Bank.

Kursa MB, Rudnicki WR (2010) Feature selection with the Boruta package. *J Stat Software* 36: 1-13.

Lempérière Y, Deremble C, Seager P, Potters M, Bouchaud JP (2014) Two centuries of trend following. *arXiv preprint arXiv:1404.3274*.

Lux T (1997) Time variation of second moments from a noise trader/infection model. *J Econ Dyn Control* 22: 1-38.

Lux T (2009a) Applications of statistical physics in finance and economics. In Jr., J. B. R. and Jr., K. L. C., editors, *Handbook of Research on Complexity*, chapter 9. Edward Elgar, Cheltenham, UK.

Lux T (2009b) Rational forecast or social opinion dynamics? Identification of interaction effects in a business climate survey. *J Econ Behav Organ* 72: 638-655.

McLeay M, Radia A, Thomas R (2014) Money in the modern economy: an introduction. *Bank of England Quarterly Bulletin* Q1.

Menkhoff L, Sarno L, Schmeling M, Schrimpf A (2012) Currency momentum strategies. *J Financ Econ* 106: 660-684.

Minsky, H. P. (1977) The financial instability hypothesis: An interpretation of Keynes and an alternative to “standard” theory. *Nebraska J Econ Bus* 16: 5-16.

Minsky HP (2008) *Stabilizing an Unstable Economy*. McGraw-Hill, New York, NY. A reissue of 1986 publication.

Moore MJ (2009) Morgan Stanley’s Mack welcomes regulation by Fed (update 2). *Bloomberg*.

Osborn DR (1995) Moving average detrending and the analysis of business cycles. *Oxford Bulletin of Economics and Statistics* 57: 547-558.

Towers G (1939) Minutes of proceedings and evidence respecting the Bank of Canada. *Committee on Banking and Commerce, Ottawa: Government Printing Bureau*.

Weidlich W (1972) The use of statistical models in sociology. *Collective Phenomena* 1: 51-59.

Weidlich W (2000) *Sociodynamics: A Systematic Approach to Mathematical Modeling in the Social Sciences*. Dover, Mineola, NY.

Weidlich W, Haag G (1983) *Concepts and Models of a Quantitative Sociology: The Dynamics of Interacting Populations*, volume 14 of *Springer Series in Synergetics*. Springer-Verlag, Berlin.

Werner RA (2014) Can banks individually create money out of nothing? – The theories and the empirical evidence. *Int Rev Financ Anal* 36: 1-19.

Werner RA (2016) A lost century in economics: Three theories of banking and the conclusive evidence. *Int Rev Financ Anal* 46: 361-379.