Ozgur *et al. Financ Innov* (2021) 7:20
https://doi.org/10.1186/s40854-021-00237-1

Financial Innovation

RESEARCH Open Access

Machine learning approach to drivers of bank lending: evidence from an emerging economy

Onder Ozgur*, Erdal Tanas Karagol and Fatih Cemil Ozbugday

*Correspondence:
oozgur@ybu.edu.tr
Department of Economics, Faculty of Political Sciences, Ankara Yıldırım Beyazıt University, Esenboğa Campus, Çubuk, Ankara, Turkey

Abstract

The study analyzes the performance of bank-specific characteristics, macroeconomic indicators, and global factors to predict the bank lending in Turkey for the period 2002Q4–2019Q2. The objective of this study is first, to clarify the possible nonlinear and nonparametric relationships between outstanding bank loans and bank-specific, macroeconomic, and global factors. Second, it aims to propose various machine learning algorithms that determine drivers of bank lending and benefits from the advantages of these techniques. The empirical findings indicate favorable evidence that the drivers of bank lending exhibit some nonlinearities. Additionally, partial dependence plots depict that numerous bank-specific characteristics and macroeconomic indicators tend to be important variables that influence bank lending behavior. The study's findings have some policy implications for bank managers, regulatory authorities, and policymakers.

Keywords: Bank lending, Machine learning techniques, Decision trees, Turkey

Introduction

There are many studies on the importance of financial intermediaries in the existing literature. Financial intermediaries help lenders to invest their wealth into activities that yield smooth returns and help borrowers to increase their real asset holdings (Brainard and Tobin 1963). Owing to problems such as asymmetric information and numerous financial frictions in financial markets, the financial intermediaries' role becomes more critical. Diamond (1984) finds that financial intermediaries are delegated monitors that specialize in minimizing potential information-based problems between borrowers and lenders. Boyd and Prescott (1986) also claim that financial intermediaries reduce the cost of collecting and processing information regarding the borrower's creditworthiness. Therefore, financial intermediaries effectively convert large amounts of savings into profitable investments (Levine 2005; Clark 2017).

The finance literature argues that banks are regarded as the most prominent financial intermediaries, among other institutions, especially in developing countries. Banks perform tasks related to risk size, maturity, and transformation by balancing borrowers' great and long-term funding needs with the collection of small and short-term savers'

Springer Open

© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

deposits (Casu et al. 2006). Banks promote and mobilize investments, increase capital performance, and make transactions easier (Seven and Yetkiner 2016). Bernanke (1993) briefly notes that particular roles are attributed to banks based on their banking officials' experiences in lending to specific industries, the banks' ability to determine the creditworthiness of potentially less risky borrowers, and the provision of various financial services to clients beyond lending. Further, Greenbaum et al. (2019) find that banks are specialized in disseminating information regarding monetary policy changes and fulfilling their role to achieve monetary stability.

Firms need internal or external support to promote their operations. As firms, especially small firms, cannot meet their financing needs entirely from internal sources, they require external funding to fulfill their goals (Imran and Nishat 2013). Asymmetric information problems in small and medium-sized firms are primarily attributed to the firms' preliminary financial reports, and they hinder the firms from acquiring bank loans to meet their financial needs (Sarath and Van Pham 2015). Gertler and Gilchrist (1993) argue that small firms and households lack various financial options and are dependent on bank loans. Since banks have a unique monitoring role, they can provide firms with a relatively less expensive financing alternative (Ramey 1993). Therefore, bank loans and other debt financing forms appear to be imperfect substitutes (Gambacorta and Rossi 2010).

Banks have a special role in all countries, with bank loans being critical in the financial and real sides of the economy, for various reasons. First, any shocks to bank loans are expected to influence the economic activity significantly (Çavuşoğlu 2002). Second, for banking institutions, bank loans are essential assets and the primary sources of the banks' income (Malede 2014). Third, the level of outstanding bank loans seems to be a vital indicator of a country's monetary stability and is related to the overall price level in any economy (Calza et al. 2003). Finally, bank loan patterns are also indicators of financial stability and are perceived as indicators of impending economic and financial crisis (Pham 2015).

After the onset of the domestic crisis in 2000 and 2001, several adjustments and regulatory measures were taken in the Turkish economy (Akinci et al. 2012). Following these measures, the Turkish economy experienced rapid credit growth and rapid increase in its economic activities in the subsequent decades (Kara 2016). The share of the banking sector is approximately 89% of the total financial system. Deposit banks are more dominant in the financial system; in the second quarter of 2019, they constituted approximately 78% of the total assets in the financial sector (EDDS 2020). Therefore, the Turkish economy offers an ideal environment to evaluate the impact of various factors on outstanding bank loans.

As bank loans are crucial for the economic activities, financial stability, and financial stance of any country, many studies have attempted to clarify the factors that influence the drivers of bank lending in both emerging and advanced economies. While more scholars are increasingly becoming aware of the factors affecting bank lending behavior, owing to the bank loans' critical role, existing studies mainly focus on econometric techniques and present mixed results (Égert et al. 2007; Brei et al. 2013; Shijaku and Kalluci 2013; Ivanovic 2016; Kapounek et al. 2017; Naceur et al. 2018). Therefore, there is scope for contribution, and the literature calls for studies on recent empirical methodologies.

The objective of this study is first, to clarify the possible nonlinear and nonparametric relationships between outstanding bank loans and bank-specific, macroeconomic, and global factors. Second, it aims to propose various machine learning algorithms that determine drivers of bank lending. Third, it aims to expand the literature by using a wide range of control variables, and fourth, it aims to provide empirical evidence regarding the role of such variables in forecasting outstanding bank loans. To this end, the study shows the impact of 19 bank-specific, macroeconomic, and global variables on bank loans for the period 2002Q4–2019Q2 in Turkey. It compares the performance of the regression model with machine learning methods, such as regression tree, boosting, bootstrap aggregating (bagging), random forest, extremely randomized trees (extra-trees), and extreme gradient boosting (xgboost). The variable specification depends on the literature, and the data set includes numerous explanatory variables that are considered to influence bank-lending behavior significantly. Policymakers should focus on monitoring the shocks to these variables to provide more empirical evidence that is relevant and based on which, variables are more important for bank lending.

Machine learning techniques are designated to use algorithms while predicting, classifying, and clustering datasets (Athey 2018). The concept of using machine learning techniques in economic analysis is a relatively novel approach (Varian 2014). These techniques can provide alternatives for existing methodologies to close the literature gaps and have several benefits. First, machine learning algorithms solve dimensionality problems in empirical studies. Conventional models can only handle a few existing studies (Fornaro and Luomaranta 2020) and, therefore, fail to benefit from large datasets (Petropoulos et al. 2019). However, this study's empirical specification allows employing 19 bank-specific, macroeconomic, and global variables to examine bank-lending drivers.

Second, machine learning techniques allow flexibility in selecting the functional form of the model (Athey 2018), without requiring any prior assumptions regarding the distribution of variables (Alessi and Detken 2018). Thus, these techniques provide the best suiting functional forms for data (Mullainathan and Spiess 2017). Third, machine learning techniques offer an out-of-sample predictive ability of variables, allowing the ranking of variables according to their out-of-sample fit measures (Basuchoudhary et al. 2017). Thus, these techniques increase both scholars' and policymakers' confidence in understanding economic problems.

The rest of this paper is structured as follows: The second section reviews the literature, and the third section presents an overview of the Turkish banking sector. The fourth section introduces the data and variables, and the fifth section explains the methodology. The sixth section presents the empirical results, and the last section discusses and concludes the study.

Literature review

Several studies have attempted to clarify the factors that significantly influence bank-lending behavior. Although the literature can be delivered in various classifications, we prefer to review the literature on developed and emerging economy studies.

Studies in developed countries focus on numerous bank-specific characteristics, macroeconomic indicators, and global factors. In an earlier attempt, Berger and Udell (1994) focused mainly on risk-based capital ratios and other capital measures to explain the

credit crunch in the United States in the early 1990s. They concluded that alternative capital ratios did not significantly affect loan growth during this period. In another study, Carlson et al. (2013) showed that capital ratios significantly influence bank lending, during and slightly after global financial crises. They also revealed some nonlinearities in the impact of capital ratios on bank lending, since the capital ratio elasticity of bank lending is somewhat higher when capital ratios are lower. In a recent study on US banks, Kim and Sohn (2017) found that the capital ratio significantly and negatively affects the loan growth for banks that have lower liquidity ratios.

In the Euro Area, Calza et al. (2003) illustrated that the impact of GDP on real loans is positive, while both short-term and long-term interest rates negatively affect real loans in the end. Kapounek et al. (2017) analyzed bank-lending drivers using a more comprehensive set of variables and found that the lending rate spread, asset quality, and capital ratios are significant bank-specific and loan supply factors. Panagopoulos and Spilliotis (1998) did not find a significant impact of interest rates on credit expansion. In Germany, Blaes (2011) demonstrated that bank lending survey variables significantly influenced the bank lending behavior during the sample period. In addition, a set of bank-related factors was confirmed to have a more substantial impact on bank lending, following the financial crisis. Similarly, Cucinelli (2015) found that credit risk was a significant factor in bank lending during the post-financial crisis, for both groups of banks in Italy.

Blundell-Wignall and Roulet (2013) focused on the impact of unconventional monetary policy settings on bank lending behavior, using an unbalanced panel of 468 US and European banks. Their findings indicate that among the total assets, the share of risk-weighted assets is a significant measure of the global systematically important banks' loan supply. Conversely, the loan supply for other banks in the sample is significantly influenced by the banks' solvency measures, the spread between lending rates, loan quality, and demand-side measures. In a recent study, Naceur et al. (2018) demonstrated that, during the post-global financial crisis, capital ratios and liquidity measures significantly influenced bank lending. They also caused heterogeneity in the characteristics of banks across Europe and the United States.

Hoffman (2001) indicated that short-term and long-term interest rates significantly influence the private sector' bank credit, and their findings, from sampled countries, also support that real estate prices have a significant effect on bank lending. By employing a sample of large international banks in 14 advanced countries, Brei et al. (2013) suggested that the capital ratio's marginal impact is more significant during the crisis than it is during regular times.

The study by Kosak et al. (2015) supports the significant impact of capital on bank lending for both periods. In a more comprehensive study, Pham (2015) studied a sample of banks from 146 countries, from 1990 to 2013. Their findings reflect the heterogeneous reactions of banks, provide evidence of the impact of country-specific factors on bank lending behavior, and emphasize a healthy financial system's role in absorbing external shocks to the banking system.

Many studies also focus on the drivers of bank lending in emerging countries. Imran and Nishat (2013) found that foreign liabilities, deposits, GDP growth, domestic monetary conditions, and exchange rates are significant long-run determinants of bank lending in Pakistan. Sarath and Van Pham (2015) revealed that the growth of deposits, equity

growth rate, liquidity, share of security holdings, and real economic growth significantly affect bank lending in Vietnam. Recently, Baoko et al. (2017) studied the factors that influence the allocation of bank lending to the private sector in the Ghanaian economy. Their findings suggest that the bank size, real bank lending rate, bank deposit rate, and broad money supply significantly influence the bank credit, both in the long and short run. Among the existing studies, Ladime et al. (2013) emphasized the role of competition in the Ghanaian economy's banking industry, and Shijaku and Kalluci (2013) focused on the impact of financial liberalization in Albania. Malede (2014) focused on internal and external factors in Ethiopia, Rababah (2015) analyzed the impact of bank-specific and macroeconomic indicators in Jordan, and finally, Ivanovic (2016) distinguished between post- and pre-crisis periods. Recent studies on the Turkish economy focus on non-performing loans and the impact of extraordinary increases in lending activities on bank riskiness. Among these studies, Us (2017) proposed that, when designing macro-prudential policies during crises, the ownership status should be considered to assess non-performing bank loans. Recently, Shahzad et al. (2019) noted that the previous year's loan growth rate had a significant impact, which resulted in the rise in loan losses and the increasing bank insolvency. Moreover, their survey noted that in the Turkish banking system, the size of banks and increase of non-performing loans are inversely linked.

Unlike studies in developing or emerging countries, machine learning studies on bank lending analysis primarily focus on risk management issues and credit ratings.[1] Among these studies, Petropoulos et al. (2019) predicted the future behavior of corporate loans and evaluated how particular measures affect the credit risk in the Greek banking system. Their findings highlight the regulatory authorities' significant role of monitoring bank-level micro data. Wang et al. (2018) argued that the hybrid xgboost model more accurately predicts credit fraud risk in banking operations. A more technical analysis by Zhu (2019) provided evidence indicating that machine learning algorithms perform better in predicting loan default. More recently, Orlova (2020) highlighted the impact of credit management models on credit institutions' profitability. For credit ratings and credit scoring, Wang et al. (2020) demonstrated the significant role of cost-sensitive classifiers for survival and profitability in peer-to-peer lending. Finally, Shen et al. (2020) provided evidence supporting the superiority of a novel three-stage learning framework in credit scoring performance by handling Chinese credit data.

Overall, the vast majority of the existing studies employs econometric techniques and provides rather mixed results in an attempt to justify the impact of bank-specific variables, macroeconomic indicators, and global factors on bank lending behavior. As such, the literature calls for studies to clarify these factors' role using a comprehensive set of variables and adopting a novel methodology, compared to that in the existing literature, and highlight the benefit or the apparent advantages of the machine learning techniques.

[1] See Leo et al. (2019) for a review of the studies focusing on the performance of machine learning techniques in bank risk management.

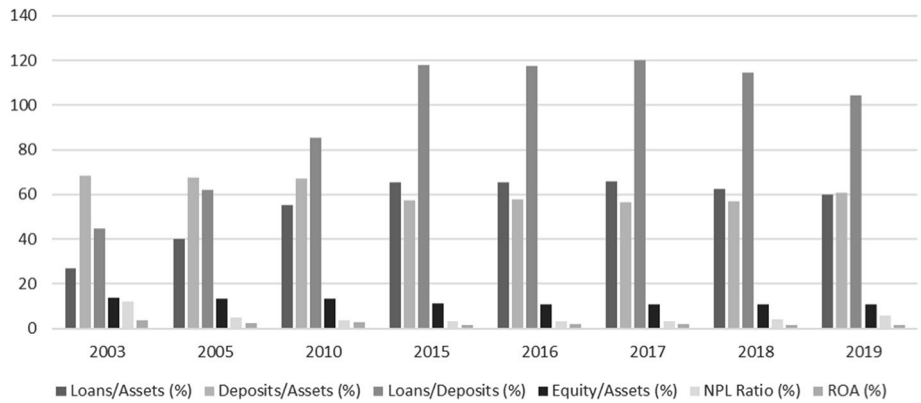

Fig. 1 Selected ratios in deposit banks in Turkey. *Source*: Banking Regulation and Supervision Agency (BRSA) Monthly Banking Sector Data

Banking sector in Turkey

Turkey's lending market comprises three main groups of banks: deposit banks, development and investment banks, and participation banks. Deposit banks are the leading participants and dominate the sector, which currently consists of 51 banks, including 32 deposit banks, 13 investment and development banks, and six participation banks. In June 2019, deposit banks provided 86.86% of the total loan supply (BRSA 2020). In contrast, development and investment banks as well as participation banks have relatively lower shares in the market and provide 8.24% and 4.89% of the total loan supply, respectively.

Figure 1 shows the structure of the deposit banking industry and highlights some selected ratios that reflect the deposit banking industry's improvements, following the two severe crises in November 2000 and February 2001.

The ratio of loans to total assets, deposits to total assets, and loans to deposits have increasingly grown over the post-crisis period, which indicates that banks are more confident in their primary duty, which is fulfilling their financial intermediary role. The capital ratio, which showed a more stable trend during this period, decreased from 13 to 11%. This indicates that banks were more focused on external financing sources than on internal sources and used more deposits to finance their lending activities and other assets. One of the most noticeable changes following the crises was that the non-performing loans (NPL) ratio improved. The NPL, which is an indicator of asset quality, decreased from 12% in 2003 to 4% in 2018. However, it increased slightly in 2019 to 5.71%. Finally, the return on asset (ROA), which indicates the banks' asset profitability, dropped from 3.5 to 1.40% in this period. As noted, a probable explanation is that, after the crises, the restructuring regulations and macroeconomic measures eliminated the more profitable non-banking activities that had emerged in the 1990s. Therefore, banks relied more on their primary functions in the financial system.

Overall, the Turkish banking sector has faced several crises, undergone many structural changes, and experienced a set of regulatory measures, particularly after introducing the financial liberalization process in the 1980s. Although the steps taken due to developments in this period are sometimes unsuccessful, the measures taken after the crises of 2000 and 2001 guided the banking sector to establish a more robust and

healthier structure. They also strengthened the role of banks in financing real economic activities.

Datasets

Data

This study's empirical analysis focuses on a sample of 19 deposit banks in Turkey that have at least ten operational branches. These banks constitute 97.13% of the total assets and 97.35% of the total loans in the deposit-banking sector. The sampled banks are listed in Additional file 1: Table A.1. The dataset contains nine bank-specific variables, seven macroeconomic indicators, and three global factors to determine the critical factors that explain bank lending behavior. Thus, the dataset employs 19 explanatory variables, which include the indicators used in previous studies to determine the factors affecting bank loans. The next section presents detailed explanations of the variables used in the analysis. The sample consists of quarterly observations and covers the period between 2002Q4 and 2019Q2; the starting date coincides with the post domestic crisis period in Turkey. Additionally, the dataset is balanced, with no missing value, and the total bank-year observations are 23,940.

The only source for the bank-level data is the statistical reports from the Banks Association of Turkey (BAT). The BAT delivers quarterly balance sheets for each bank settled in Turkey. Other sources for the wide-range dataset are the Thomson Reuters Eikon and Electronic Data Delivery System, which are provided by the Central Bank of the Republic of Turkey. The full list of these variables, their acronyms, descriptions, units, and sources, is shown in Additional file 1: Table A.2.

Variables

This study's model specification to estimate the drivers of bank lending in Turkey employs the natural logarithm of nominal bank loans (Ln(Loans)) as a measure of the extended bank loans. This variable includes consumer and commercial loans, in both domestic currency and foreign currency. However, all figures are in Turkish Lira (TL) since banks' balance sheets include foreign-currency-dominated loans in TL terms. The model embodies a comprehensive set of variables, as explanatory variables, in three groups, including bank-specific variables, macroeconomic indicators, and global factors. These selected variables h play a significant role in banks' lending behavior, as indicated in existing studies (see Additional file 1: Table A.2 for variable definitions and groupings).

Since bank lending is not contemporaneously affected by the changes in bank-specific characteristics, macroeconomic indicators, or global factors, we included each variable's highly correlated lag into the analysis.[2] The correlations presented in Additional file 1: Table A.3, either include the current or lagged values of the independent variables into the analysis. To illustrate, the third lag of the ROA has the highest correlation, while the fourth lag of the LLPTA has the highest correlation, with an independent variable. Most of the explanatory variables were lagged to mitigate the unintended feedback effects due

[2] Additional file 1: Table A.3, presents the correlations between the natural logarithm of nominal bank loans and the set of employed variables, up to their fourth lags.

Table 1 Descriptive statistics

Group	Acronym	Lag	Unit	Mean	Median	Max	Min	Std. Dev
Dep. Var	Ln(Loans)	Current	TL	16.005	16.175	19.665	3.611	2.064
Bank-Specific Characteristics	LLPTA	4	Ratio	0.017	0.015	0.085	0.000	0.012
	LIQR	Current	Ratio	0.253	0.241	0.781	0.006	0.121
	GVTTA	Current	Ratio	0.172	0.144	0.774	0.000	0.128
	FATA	4	Ratio	0.020	0.013	0.516	0.001	0.030
	DFR	Current	Ratio	0.629	0.621	0.875	0.012	0.091
	DEPOVER	4	Ratio	− 0.082	− 0.020	1.640	− 0.984	0.310
	CAPR	1	Ratio	0.121	0.114	0.916	0.033	0.051
	ROA	3	Ratio	0.008	0.008	0.059	− 0.176	0.014
	LENDINGR	4	Ratio	0.083	0.074	1.589	0.018	0.065
Macroeconomic Indicators	Ln(GDP)	4	Thousand TL	19.554	19.487	20.609	18.383	0.595
	LEAD	Current	Index	152.992	139.380	215.910	120.770	30.806
	CPI	4	Index	106.766	99.731	195.931	48.800	38.714
	ON	4	%	12.659	10.500	44.000	1.500	8.885
	REER	Current	Index	101.348	101.355	118.183	76.655	9.351
	GVT9M	4	%	14.617	11.001	54.700	4.460	10.576
	GVT2Y	4	%	24.407	11.300	81.690	5.670	26.726
Global Factors	FFR	Current	%	1.339	0.400	5.260	0.070	1.662
	ECB	4	%	1.467	1.000	4.250	0.000	1.301
	OIL	4	US Dollar	71.114	66.725	121.477	26.067	28.389

to endogeneity issues (Berger and Udell 1994). The list of current or lagged values of the employed variables and their descriptive statistics, resulting from this specification, is presented in Table 1.

The dependent variable in our specification is the natural logarithm of the nominal bank loans. This study considers the natural logarithm of outstanding bank loans to reduce outliers' impact, since it covers 19 deposit banks in various sizes (Levine 2005). Nominal bank loans are the outstanding amount of loans in households and firms, for a sample comprising 19 deposit banks. This measure of bank loans has been used in existing studies (Panagopoulos and Spilliotis 1998; Gambacorta and Rossi 2010). Some studies have used a similar measure as a dependent variable (Hoffman 2001; Calza et al. 2003). However, our analysis differs from these studies in that it integrates potential non-linearities into the model and benefits from the machine learning algorithms to provide the best predictors for outstanding bank loans.

The visual representation of the relationship between outstanding bank loans and the set of employed variables is shown in Fig. 2.

The first set of explanatory variables comprises the bank-specific characteristics, derived from deposit banks' balance sheets and income statements. These variables are used to assess how bank characteristics affect bank lending. Some studies suggest that bank lending is heterogeneous among banks with different characteristics (Kashyap and Stein 1995).

To control for the bank lending behavior, this study introduces a set of macroeconomic indicators into the model, which include Natural logarithm of Gross Domestic Product

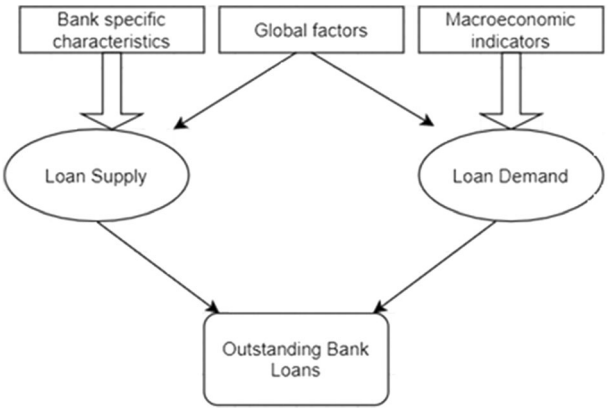

Fig. 2 Interaction between bank loans and predictor variables

((Ln(GDP)), Leading Indicators (LEAD), Consumer Price Index (CPI), Overnight Interest Rate (ON), 9-month Treasury Bill rate (GVT9M), 2-year Government Bond rate (GVT2Y), and the Real Effective Exchange Rate (REER) series. The GDP and LEAD are included to highlight the procyclical behavior of bank lending; CPI to analyze the impact of price stability on bank lending; and interest rate variables to analyze the role of monetary policy rate and market rates on outstanding bank loans. The REER explores the role of the exchange rate on bank lending behavior. A rise in the REER represents the appreciation of the domestic currency in our specification.

Other than bank-specific characteristics and macroeconomic indicators, this study also examines the role of global factors on bank lending. Since Turkey is a small open economy, global financial and economic development might affect the domestic economy (Varlik and Berument 2017). As such, this study employs the US Federal Funds Rate (FFR), European Central Bank (ECB) main refinancing rate (ECB), and the crude oil price (OIL). Of these variables, FFR and ECB specifically account for the impact of global liquidity (Pham 2015) and reflect the cost of foreign banks' borrowing (Ivanovic 2016) on outstanding bank loans. Finally, the OIL variable is included in the analysis to capture global developments (Potjagailo 2017). Changes in oil prices might have adverse effects on investment and consumption expenditures, increase production costs, and reduce firms' cash flows (Kocaarslan and Soytas 2019). Therefore, the impact of oil prices on bank lending behavior should be regulated.

Before proceeding into with the analysis, this study first illustrates the relationship between employed variables and outstanding bank loans. Figure 3 illustrates the relationship between predictor variables and outstanding bank loans.

Figure 3 demonstrates the potential nonlinearities between predictor variables and outstanding bank loans. In this case, specifying models that disregard these nonlinear linkages might create misleading outcomes. Therefore, these figures support the use of machine learning algorithms.

Methodology

This study extends the methodology proposed by Basuchoudhary et al. (2017). It compares the performance of six machine learning techniques (tree regression, bagging, boosting, random forest, extra-trees, and xgboost) and the standard linear regression

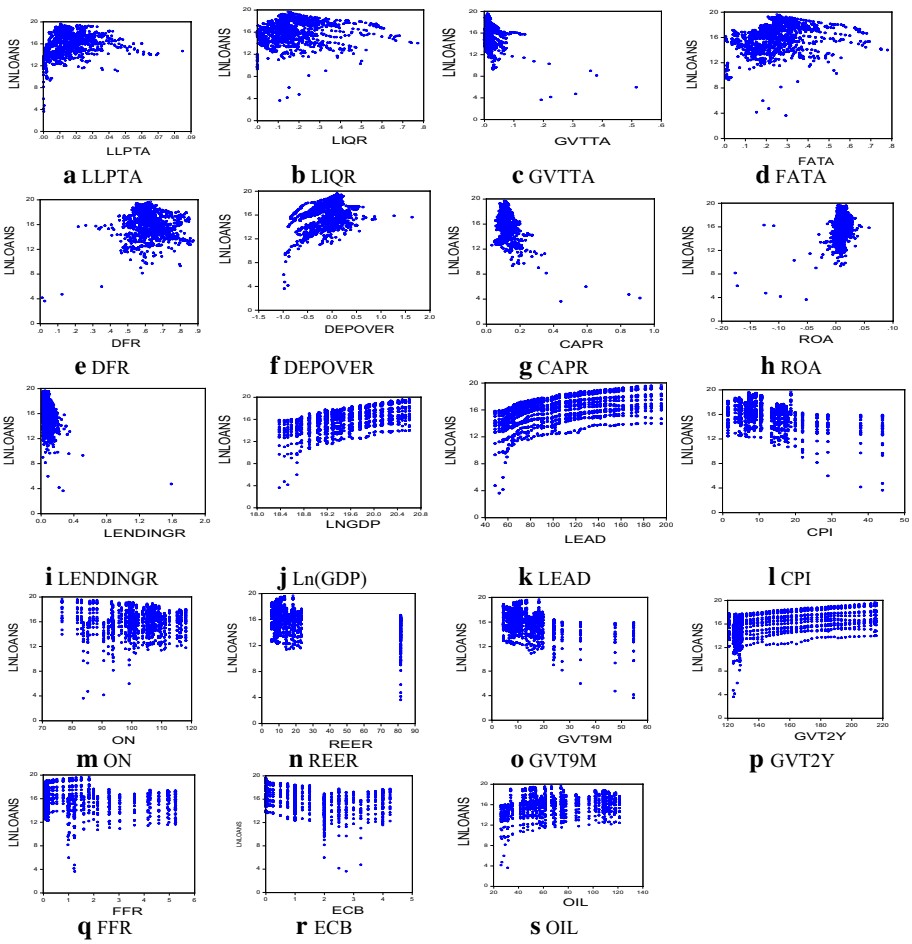

Fig. 3 Relationship between bank loans and predictor variables

to predict factors that influence commercial bank lending operations in Turkey. The properties of such techniques should be addressed before introducing the empirical specification of the study. The purpose of the machine learning techniques is to predict the output variable (Y), for an independent test sample, using the learning sample, where both input (X) and output variables are observed (Athey 2018).

The standard linear regression models represent the parametric models, and they assume a linear functional form $Y = f(x)$. These models are generally easy to implement, the estimated coefficients are easy to interpret, and deciding the statistical significance of these coefficients is easy (James et al. 2013). However, these models generally cannot handle large dimensional datasets (Fornaro and Luomaranta 2020). In other words, they only estimate a small number of variables in any specification. On the contrary, if the linearity assumption is not in place, it is reasonable to assume that these models do not fit the data best.

Unlike the standard techniques, the machine learning techniques allow researchers to handle a sizeable dimensional dataset (Fornaro and Luomaranta 2020), provide model flexibility without forcibly adjusting the specification to a specific functional

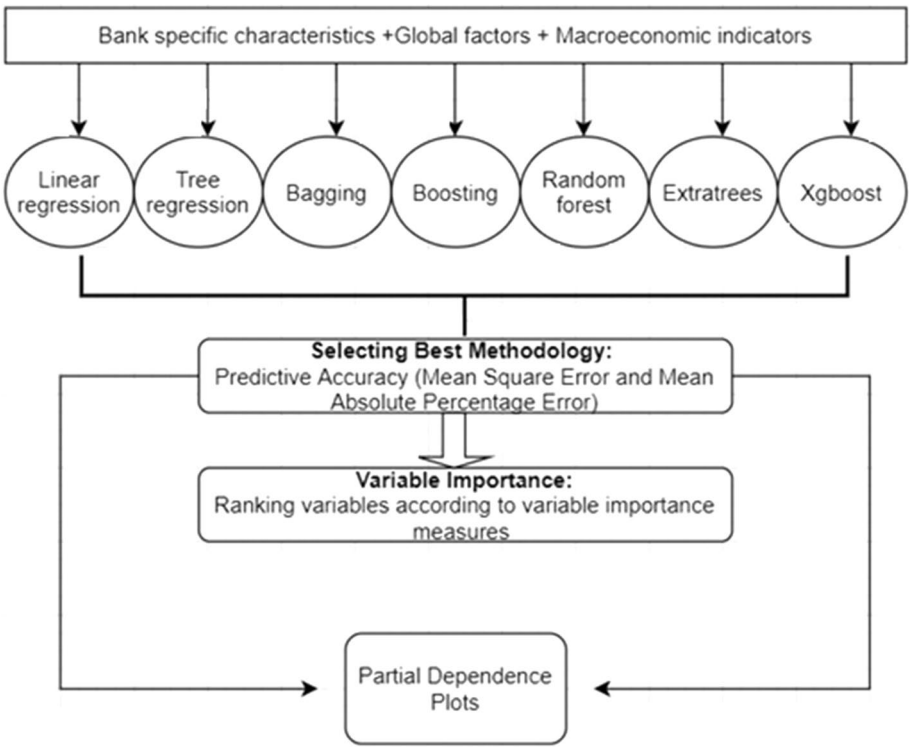

Fig. 4 Visual representation of empirical strategy

form (Mullainathan and Spiess 2017), and do not require any prior assumptions regarding the distribution of variables (Alessi and Detken 2018).

Before discussing each machine learning algorithm's details, we summarize the preliminary steps to run these algorithms.

- First, we collect the observations on bank-specific characteristics, macroeconomic indicators, and global factors and subsequently organize the dataset.
- Since machine learning algorithms repeatedly resample within the learning sample, in an attempt to find the one that fits the data best (Basuchoudhary et al. 2017), we divide the dataset into two random subsets of banks: a learning sample and a test sample. The randomly chosen 70% of the banking data is used as the learning sample and the remaining part (test sample) is for the out-of-sample validation purpose.
- We do not use cross-validation but handle an entirely separate test sample instead.
- After these steps, we run the machine learning algorithms together with the linear regression and compare their predictive accuracy to find the model that best fits the out-of-sample (test sample) data.

The empirical part of the study is structured as shown in Fig. 4.

First, we run a linear regression, and the machine learning algorithms handle our dataset's learning sample. Subsequently, the Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE) are selected and then calculated as a measure of these empirical models' out-of-sample predictive ability. Next, the variable

importance measures are assessed and the bank characteristics, macroeconomic indicators, and global factors, ranked based on their role in predicting the bank-lending behavior. Finally, this study presents partial dependence plots (PDPs) to illustrate the direction, magnitude, and impact of input variables on bank lending behavior. The following sections present the theoretical illustration of the employed machine learning techniques.

Regression tree prediction

Unlike linear models, a regression tree prediction aims to divide a sample into subsamples, based on the impurity of the target variable (y_i), using if–then statements as parameters (Basuchoudhary et al. 2017). This study follows these steps in the regression tree algorithm (Breiman et al. 1984):

1. The learning sample is divided into terminal nodes (splitting variable) and internal nodes (splitting point), using a decision rule $\hat{d}(x)$.
2. The decision rule determines the splitting points for each node by minimizing the impurity within the nodes. It also includes an externally determined, threshold-level impurity change measure and the minimum number of observation requirements within each node (Basuchoudhary et al. 2017).
3. Therefore, the best terminal node and the best internal node produce a higher node impurity change.
4. The procedure reaches a decision node when splitting is not possible with given decision rule requirements.

Regression trees have many advantages. They are easy to interpret and explain, and they can include qualitative variables in the analysis without the need for dummy variables (James et al. 2013).

Boosting prediction

The boosting algorithm is a technique used to improve the prediction, and it is based on the decision tree. The boosting methodology was first proposed by Freund and Schapire (1997). It was initially developed for classification purposes and benefited from regressions (Hastie et al. 2017).

The boosting algorithm in our specification follows these steps (James et al. 2013):

1. The boosting algorithm constructs M different learning samples and assigns a weight to each test sample.
2. The algorithm then fits each of these M learning samples into different decision trees
3. These decision trees provide the predictions as $\widehat{f^1}(x), \widehat{f^2}(x), \ldots, \widehat{f^M}(x)$.
4. The error rates for each of these predictions are calculated. New weights are assigned if the error rate is reasonable, based on the importance of the variable.
5. The algorithm, therefore, slowly improves $\hat{f}(x)$ performance when required, by fitting small trees to the residuals.

Bagging (bootstrap aggregating) prediction

The bagging predictor was first proposed by Breiman (1996). This algorithm uses portions of data, creates independent predictors, and then combines each by averaging to get an aggregated predictor. The experimental procedure of the bagging predictor in this study is as follows (Breiman 1996):

1. The algorithm takes M bootstrapped sub-samples from the learning data, in which each subsample includes N independent observations.
2. M decision trees are constructed using these M bootstrapped sub-samples.
3. The algorithm provides M decision trees, $\widehat{f^1}(x), \widehat{f^2}(x), \ldots, \widehat{f^M}(x)$, from M tree predictions.
4. The bagging predictor is the average of the sequence of tree predictions, and it is illustrated as $\widehat{f^{ave}}(x)$; the total error is the average of the errors in each tree's predictions.

Notably, the bagging methodology forms the learning sample by randomizing observations with replacement, and it defines splits for each node by minimizing the node impurity to grow M trees. The bagged tree is the average of the predictions from M trees (Hastie et al. 2017).

Random forest prediction

The random forest algorithm is another tree-based algorithm used for regression purposes. It was first proposed by Breiman (2001). The random forest predictor randomizes the trees by selecting bootstrapped observations as learning samples and selecting input variables for each tree (Basuchoudhary et al. 2017).

Empirical steps for the random forest algorithm are as follow (Breiman 2001).

1. It first, independently, draws a bootstrap sample of size N from the learning sample.
2. Afterwards, it randomly selects k variables out of the total p ($k \leq p$), and these variables are possible candidates for splitting in each node.
3. In the next step, the M random forest trees are grown repeatedly until the minimum node size is reached.
4. The random forest regression predictor is the average of these M trees.

Extremely randomized trees (extra-trees) algorithm

The extra-trees algorithm is a tree-based ensemble algorithm that uses a conventional top-down procedure (Geurts et al. 2006). The empirical procedure of this algorithm is quite similar to that of the random forest algorithm (Altabrawee 2017). There are two main differences between the extra-trees algorithm and the random forest algorithm (Geurts et al. 2006):

1. The extra-trees algorithm randomly chooses the splitting points at each node.

2. The random forest algorithm employs the entire learning sample while growing trees, instead of bootstrapping a sample of size N from the learning sample.

Geurts et al. (2006) argue that the randomness in the training (learning) part of the procedure increases independent trees and decreases the variance.

Extreme gradient boosting (Xgboost) algorithm

Extreme gradient boosting or Xgboost is an extended version of Friedman's (2001) gradient boosting algorithm proposed by Chen and Guestrin (2016). Xgboost algorithm uses the bagging method to reduce bias, the boosting method to decrease variance, and fits regressions to increase efficiency and accuracy in objective functions (Petropoulos et al. 2019).

The algorithm in our study follows these steps (Tang et al. 2020):

1. The Xgboost algorithm creates M bootstrapped samples of the learning data and creates M decision trees using these sub-samples.
2. These decision trees provide a tree prediction and bags these M tree predictions. Note that each tree transmits the subsequent tree's errors to reduce the existing errors in each tree.
3. The Xgboost prediction is the sum score of M boosted tree predictions.

Extreme gradient boosting offers a faster learning process among its counterparts, using a parallel tree boosting to create bootstrapped trees (Zhu 2019). It aims to increase efficiency by boosting parameters to perform the tree predictions (Petropoulos et al. 2019).

Predictive accuracy, variable importance, and partial dependence plots

The study provides two commonly used predictive indicators, MSE and MAPE, to evaluate the prediction performance of linear panel regression and six machine learning algorithms. The MSE and MAPE are denoted as $R_{MSE}(d)$, and $R_{MAPE}(d)$, respectively. The following equations present the formula for the MSE and MAPE measures.

$$R_{MSE}(d) = \frac{1}{n}\sum_{i=1}^{N}(y_i - \hat{f}(x_i))^2 \text{ and } R_{MAPE}(d) = \frac{1}{n}\sum_{i=1}^{N}\left|\left(y_i - \hat{f}(x_i)\right)/\hat{f}(x_i))\right| * 100\% \tag{1}$$

Here, the $R(d)$ is the mean of the squared difference between the target variable y_i and the predicted value of the target variable is $\hat{f}(x_i)$. Both measures provide these models' prediction performance, based on the actual and predicted values of a dependent variable.

The machine learning algorithms utilize a variable importance measure to provide the input variables' relative predictive power (Hastie et al. 2017). A specific variable's importance is calculated by summing the impurity reduction, when the specific variable is selected in splits (Tuffery 2011). Specifically, the error reduction at each split in each tree is the importance measure of the splitting variable and is measured separately for each variable, for all the trees (Hastie et al. 2017).

Table 2 Prediction performance of empirical methods in test sample

Methods	MSE	MAPE	R-squared
Panel OLS Regression	1.928	0.067	0.385
Tree Regression	3.292	0.082	0.430
Bagging	2.091	0.065	0.504
Boosting	1.732	0.061	0.588
Random Forest	**1.534**	**0.061**	**0.617**
Xgboost	1.715	1.911	0.611
Extra-Trees	1.753	0.067	0.566

Bold values represent that random forest model provides the best out-of-sample performance

Finally, the study presents PDPs to display the functional form and direction of the relationship between these input variables and bank loans. A PDP is also an illustrative measure for visualizing the relationship between input and output variables.

Suppose that x_1 is a predictor variable with values $\{x_{11}, x_{12}, x_{13}, \ldots, x_{1k}\}$. The partial dependence of the predictor x_1 can be constructed as follows (Greenwell 2017):

1. For all values of the predictor x_1, the algorithm copies the learning sample and replaces the actual values of x_1 with a constant of x_{1i}.
2. Subsequently, the algorithm calculates the predicted value vector from the modified copy of the learning sample in Step 1.
3. The algorithm then calculates the average prediction and gets $\widehat{\bar{f}}(x_{1i})$.
4. Finally, the PDP is constructed by the plotted pairs of $\left\{x_{1i}, \widehat{\bar{f}}(x_{1i})\right\}$.

Therefore, the PDPs demonstrate the marginal impact of a selected input variable on the target variable after holding the impact of other variables constant (James et al. 2013).

Empirical results

As noted, this study's empirical specification is an extended version of the general set up by Basuchoudhary et al. (2017). The model specification comprises panel OLS regression and the following six machine learning algorithms: tree regression, boosting, bagging, random forest, extra-trees, and xgboost algorithms. The subsequent sections present the predictive performance of these empirical models, provide relative importance of variables, and illustrate PDPs.

Evaluation of the models by their prediction performance

As noted, the employed algorithms' prediction performance is measured using the MSE and MAPE in our study. Both measures provide the out of sample error rates (in Eq. 1 above) and indicate which model performs better than the others do in predicting the bank loans in our specification. MSE and MAPE are commonly employed as predictive indicators to compare alternative models' prediction performance (Alpaydin 2014; Tang et al. 2020).

Table 3 The rank of variable importance in the random forest algorithm

Rank	Variable	%IncMSE	IncNodePurity	Average
1	CAPR	11.621	14.633	13.127
2	DEPOVER	10.695	9.422	10.059
3	LLPTA	8.476	8.832	8.654
4	CPI	8.070	6.506	7.288
5	LIQR	7.279	7.219	7.249
6	DFR	7.184	6.770	6.977
7	ROA	6.059	7.679	6.869
8	FATA	6.242	6.898	6.570
9	GDP	6.910	5.572	6.241
10	LEAD	6.708	5.160	5.934
11	GVTTA	5.916	5.909	5.912
12	ECB	3.866	3.209	3.538
13	GVT2Y	2.583	2.347	2.465
14	ON	2.436	2.463	2.450
15	GVT9M	2.171	2.373	2.272
16	OIL	1.382	1.285	1.334
17	LENDINGR	0.502	1.590	1.046
18	FFR	1.006	1.064	1.035
19	REER	0.897	1.068	0.982

The ranking is based on the average, and numbers are in %

Table 2 reports the MSE and MAPE measures of the panel OLS regression model and six machine learning algorithms.

The results in Table 2 indicate that not many differences exist in the predictive performance of the employed algorithms. Among these specifications, the random forest provides the lowest MSE and MAPE, outperforming its counterparts. The random forest algorithm's MSE and MAPE in the test sample are 1.534 and 0.061, respectively. Importantly, the panel OLS regression compares machine learning algorithms' performance with that for conventional econometric techniques. R-squared measure also provides evidence that the random forest model fits the sample better. It seems that an econometric model outperforms some of these algorithms. However, more tree algorithms perform better than the econometric specification in our study. Among these algorithms, the tree regression performance is worse than its counterparts are, and it provides a relatively higher MSE for out-of-sample prediction.

Variable importance

After confirming that the random forest algorithm outperforms its counterparts and provides the lowest out-of-sample prediction errors, our empirical strategy evaluates each variable's importance. To identify the most important variables, the random forest algorithm offers important measures, such as (1) mean decrease in accuracy (%IncMSE) and (2) mean decrease in node impurity (IncNodePurity). To rank these variables further, we report the average of these two measures. Table 3 reports the rank of the 19 variables, in terms of their importance in predicting outstanding bank loans.

The mean decrease in accuracy measure is the out of sample prediction error (MSE in each tree) of each tree and is calculated after permuting each variable. Therefore, Table 3

shows the percentage increase in Model-MSE if a variable for each algorithm is removed from the model. To illustrate, if CAPR is removed from the model, the MSE increases by 12.51% in the tree model, 11.62% in the random forest model, 27.54% in the bagging model, and 21.04% in the boosting model. The total is 100% in each column. The overall decrease of node impurities, due to splitting on the variable averaged over all trees, is another measure. It is determined by the residual sum of squares in a regression setting.

The results in Table 3 indicate that most of the bank-specific characteristics and macroeconomic indicators have relatively more predictive power than the global factors do. These findings seem to suggest that capital ratio, deposit overhang, and credit risk measures are the most powerful predictors. Further, CPI, GDP, and LEAD have more predictive power, among other macroeconomic indicators. Although its relative importance is lower among the input variables, the ECB main refinancing rate seems to be the most influential global factor.

Partial dependence plots

As with the most predictive power, random forest algorithm results provide meaningful policy implications for policymakers. Unlike many other empirical methodologies, the random forest algorithm is not parameter-based. Instead, they offer PDPs. As noted, PDPs illustrates the direction, impact, and functional form of the relationship between the input variable and output variable, holding others' impact constant. The distinct benefit of the random forest methodology through the PDPs' illustration is that they illustrate the incremental effect of the selected variables over the range of a specific variable. The PDP uses predicted Ln(Loans) values to construct PDPs and visualize the marginal effect of predictor variables, as shown in the PDP's empirical steps above.

In this context, the following figures display PDPs, from which the incremental impacts of bank-specific variables, macroeconomic indicators, and global factors on the outstanding bank loans can be visualized. In each PDP, the vertical axis represents the volume of bank loans in the natural logarithm, and the horizontal axis displays the range of input variables. Figure 5 illustrates the PDPs for bank-specific characteristics.

The PDP for CAPR, shown in panel a of Fig. 5, suggests that the capital ratio's effect on bank loans is not linear over the capital ratio spectrum. The figure shows that the capital ratio's positive impact on bank loans increases slightly until the capital ratio hits just over 10%. This finding is consistent with studies that provide a positive relationship between capital ratio and bank lending (Berrospide and Rochelle 2010; Cantero-Saiz et al. 2014; Ivanovic 2016; Kapounek et al. 2017). Such studies suggest that bank capital is vital against adverse shocks to bank soundness and supports banks in overcoming the negative consequences of these shocks.

When the capital ratio is marginally above 10%, even a small increase in the capital ratio will produce a more substantial adverse effect on bank loans, suggesting that after 10%, a small increase in CAPR would result in an equilibrium shift. In particular, a capital ratio that rises to 20% has a negative impact on bank lending. These results provide evidence that regulatory capital plays a conservative role in bank lending (Cucinelli 2015; Pham 2015). The negative association between bank capital and bank lending in Turkey is also consistent with the findings by Aktaş and Taş (2007) and Macit (2012). After 20%, further capital ratio improvements have a minor and smoother effect on bank

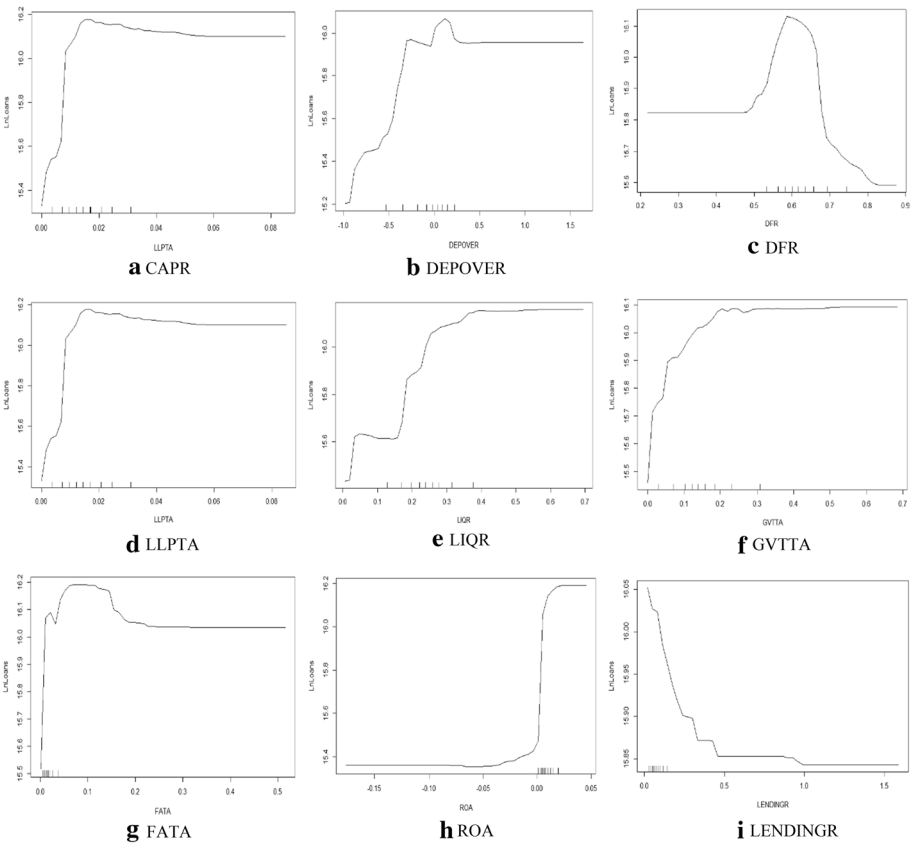

Fig. 5 Partial dependence plots for bank specific characteristics

loans. Some studies obtained somewhat mixed results and indicated that bank capital plays an insignificant role in bank loans (Bertay et al. 2012; Rabab'ah 2015).

The PDP for the effects of deposit overhang on bank loans is shown in Fig. 5b. As noted, a bank is in a deposit overhang when deposits are larger than the bank loans. To illustrate, when bank loans and deposits are equal, the DEPOVER measure is equal to zero. Figure 5b shows that bank lending is positively affected by deposit overhang. The impact of deposit overhang measure seems to increase steadily up to the point where a balance between deposits and loans in the bank balance sheet is achieved. It should also be noted that, if bank loans exceed bank deposits, the relationship between deposit overhang and bank loans reaches an equilibrium. These findings are consistent with Brinkmeyer's (2014) findings and suggest the critical role of insured deposits in financing lending activities.

Figure 5c further illustrates the relationship between the share of deposit financing and bank loans. Although the predictive salience is relatively lower than DEPOVER, similar evidence is provided about the impact of the deposit-funding ratio on outstanding bank loans.

The PDP for DFR further supports the nonlinear linkage between deposits in total assets and bank lending. It seems that the impact of DFR on bank loans reaches an

equilibrium when a balance between the share of deposits and the non-deposit source of liabilities is achieved. These findings are similar to those by Bertay et al. (2012) or Rabab'ah (2015). Such studies suggest a positive, but insignificant, coefficient of deposit funding ratio.

However, the small rise in the share of deposits causes a relatively higher increase in bank lending after the 50% threshold level. This finding suggests that the optimal share of deposit financing is approximately 60%. This specific range of the DFR somewhat supports the proposition that banks with relatively higher deposit shares, than they do other liabilities, increase their lending. In this sense, our findings are partially consistent with previous studies (Gambacorta and Marquez-Ibanez 2011; Sarath and Van Pham 2015; Malede 2014; Ivanovic 2016) arguing that deposit financing stabilizes the adverse effects of financial downturns. However, Fig. 5c indicates that the subsequent rise in the deposit shares contributes to a decrease in bank loans. Therefore, our findings suggest that reliance on deposit financing is somewhat contractionary.

The loan loss provisions to total assets, as a measure of credit risk, also have relatively higher predictive power on bank lending behavior. As shown in Fig. 5d, the LLPTA has a positive impact on bank loans when credit risk measure is relatively low. Surprisingly, it seems that the positive marginal impact steadily increases until the LLPTA approaches 2%. A possible explanation for this might be that banks remain silent to credit risk measures up to a threshold level, preferring to increase their lending.

However, after the credit risk measure falls slightly below the 2% level, the LLPTA impact alternates and turns negative. Our findings here confirm that worsening credit quality reduces bank lending. Therefore, it may be inferred that the rise in the credit risk measure puts pressure on banks' capital and causes banks to limit their lending (Naceur et al. 2018). These results are partially in accord with those of previous studies (Cantero-Saiz et al. 2014; Sarath and Van Pham 2015; Cucinelli 2015; Kim and Sohn 2017; Naceur et al. 2018), which indicate that banks with higher credit risk tend to curtail their loans.

Figure 5e illustrates the relationship between the ratio of liquid assets to total assets and outstanding bank loans. The PDP reflects that banks with higher balance sheet liquidity tend to issue more loans over a specific range of LIQR values. This suggests that banks hold liquid assets to stimulate their lending activities (Naceur et al. 2018). Another possible inference from a gradually rising liquidity ratio is that banks compensate for the cost of holding liquidity by granting more loans. The secular increase in bank loans' marginal impact on outstanding bank loans becomes stable when the liquidity ratio is higher than 40%. The PDP reflects that the association between liquidity and bank lending reaches an equilibrium after this specific point.

In conclusion, these results are consistent with those of previous studies (Imran and Nishat 2013; Malede 2014; Demiralp et al. 2017; Kim and Sohn 2017) and suggest that banks with higher liquidity can benefit from having more liquid balance sheet to absorb the impact of adverse shocks on their loan supply. The literature also includes some studies that obtain somewhat mixed results on the impact of liquidity in lending behavior (Akinci et al. 2012; Sarath and Van Pham 2015; Cantero-Saiz et al. 2014, Rabab'ah 2015), which are contrary to our conclusion.

The ratio of government securities over total assets is an alternative measure of balance sheet liquidity in our specification. The PDP for government securities over total

assets, in Fig. 5f, suggests that a rise in the share of government security portfolio, up to 20%, increases bank lending gradually. In other words, banks that have a large number of government securities in their asset portfolio increase their lending more appropriately. After the government security portfolio exceeds 20%, GVTTA and bank loans' relationship reaches an equilibrium.

Our findings seem to be in line with those by Çavuşoğlu (2002) as well as Berrospide and Rochelle (2010). Thus, these results first argue that the government security portfolio triggers bank lending since banks use their security stock as collateral to borrow from the central bank. The positive impact of government security portfolios on bank lending is called the crowding-in-effect (Çavuşoğlu 2002). This means that, when banks hold government securities in their asset portfolio, they do not cut back their private lending. Second, government securities seem to play against the possible adverse effects of the banking system's deposit drains.

A bank balance sheet's asset structure seems to be a more salient predictive measure of bank loans, compared to the bank's share of government securities. In our description, a bank's asset structure is indicated by its share of fixed assets over total assets. Figure 5g displays the PDP for the ratio of fixed assets to total assets. As noted, fixed assets comprise bank buildings, equipment, networks, intellectual properties, and other tangible and intangible assets. All these assets facilitate the financial intermediation task of banks and ease their credit-granting operations. Thus, a positive relationship is expected between the share of fixed assets and bank lending (Kosak et al. 2015). Our results, therefore, are in line with prior expectations. The PDP suggests that the small increase in fixed assets ratio to total assets leads to a substantial increase in bank loans. However, the outstanding relationship between the share of fixed assets and bank loans reaches an equilibrium when the share of fixed assets arrives at 20% of total assets.

The bank profitability is also a probable measure of bank balance sheet strength, and it might directly affect several soundness indicators (Kim and Sohn 2017). The potential positive relationship between bank lending and bank profitability is likely to be related to bank profits' role that serve as a buffer to absorb the shocks. Therefore, we conclude that the link between bank profitability measure and bank lending is in equilibrium. As long as banks gain positive profits, short-term losses do not affect bank loans. The range of negative values for ROA provides evidence for the insignificant impact of bank-lending behavior's profitability.

However, outstanding bank loans considerably increase when banks experience positive profits. Thus, it seems that the break-even point is a certain threshold level, and positive profitability causes an equilibrium shift on an outstanding relationship. As the income generated from their operations exceeds these operations' cost, banks briefly increase their lending. These findings are similar to those that indicate a positive and significant impact of profitability on bank lending (Brissimis et al. 2014; Jimenez et al. 2012; Naceur et al. 2018). Following a small increase in profitability in a positive range provides an equilibrium between ROA and bank loans. The findings that support the insignificant impact of a rise in profitability on bank lending behavior are consistent with the findings by Pham (2015) or Ivanovic (2016). The probable theoretical explanation for this finding is based on two opposing forces. First, the

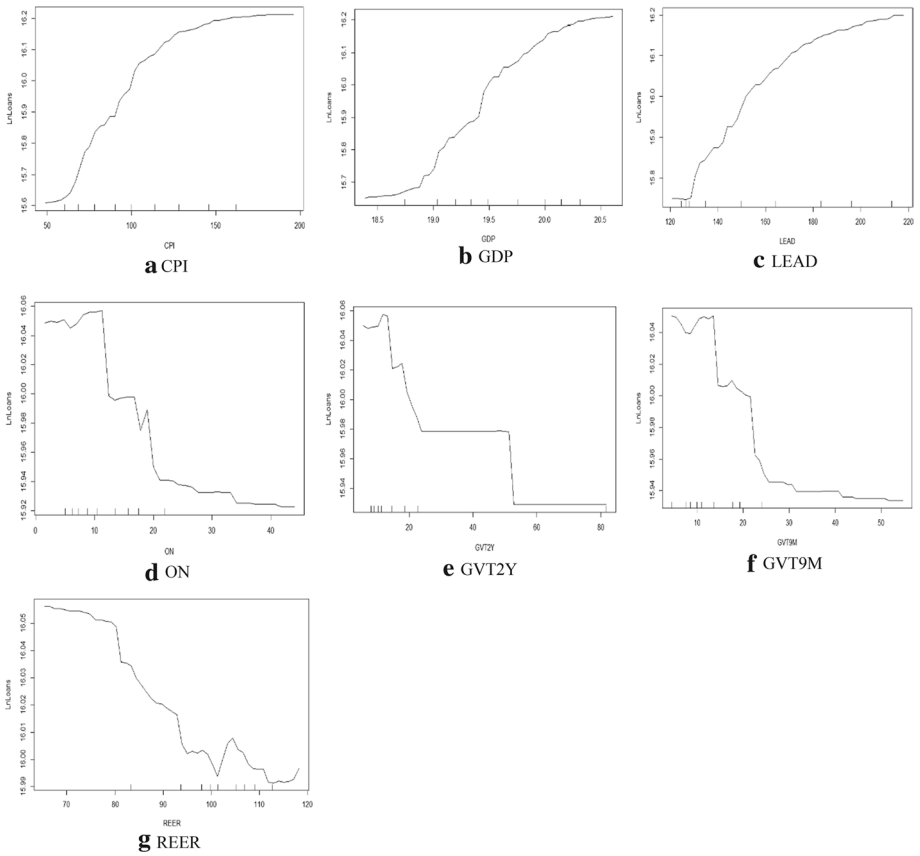

Fig. 6 Partial dependence plots for macroeconomic indicators

potential link between profitability and riskier asset portfolio directs banks to lower their lending thus improve their asset quality. Second, profitability is a signal for the sound balance sheet and encourages banks to increase their lending (Kim and Sohn 2017). Thus, these two forces might be in balance, and there might be no impact of profitability on bank lending behavior.

Surprisingly, the lending rate seems to be one of the less salient predicting measures in our random forest model. The PDP for LENDINGR suggests that the lending rate and bank loans' relationship steadily decreases until the lending rate measure is 50%. The PDP also provides evidence that when the lending ratio measure exceeds 50%, an equilibrium prevails between lending and bank loans.

The two opposing forces interact when the lending rate decreases. First, the lower lending rates stimulate borrowers' and household's loan demands (Rabab'ah 2015). Second, the lower rates discourage banks from decreasing their loan supply (Pham 2015). Thus, the former outweighs the latter, and a rise in the lending rate increases bank lending. The negative relationship between the lending rate and bank lending is similar to the findings by Gambacorta and Rossi (2010), Demiralp et al. (2017) and Baoko et al. (2017).

Among macroeconomic indicators, CPI is the most potent variable in predicting bank lending behavior. Figure 6a shows the PDP that indicates the marginal impact

of CPI on nominal bank loans. The figure clearly illustrates that there is a secularly increasing relationship between CPI and nominal bank loans. Our results, therefore, do not support the theoretical proposition suggesting that inflation has a critical role to create distortions in financial markets. In this sense, contrary to the findings in several previous studies (Panagopoulos and Spilliotis 1998; Égert et al. 2007; Jimenez et al. 2012), our results imply that inflation tends to accelerate the volume of nominal bank loans. The one possible explanation for this result is that nominal bank loans might adopt the valuation effect of a rise in the general price level. The literature also suggests that nominal bank loans increase in inflationary periods, due to increased loan demand (Kapounek et al. 2017). One might also argue that since money in hand is costly during inflationary periods, households and investors invest in deposits, which triggers the money creation process (Baoko et al. 2017). The positive impact of Turkey's CPI is in line with the literature (Alper et al. 2012).

Figure 6b and with c, comprehensively, show whether bank lending is procyclical and/or demand-driven. These figures illustrate the impact of both GDP and LEAD on bank loans. The PDP for GDP indicates a gradually increasing relationship between economic activity and outstanding bank loans. Therefore, our results might suggest that banks benefit from favorable economic conditions and improve their credit facilities. As noted, GDP is also a proxy for demand conditions in an economy. Thus, the PDP for GDP provides evidence for the critical role of demand-driven factors on the volume of bank loans. More precisely, a rise in GDP improves economic agents' earnings, creates investment opportunities, and enables the bank to lend more to finance these investments (Imran and Nishat 2013).

Figure 6c indicates the relationship between the LEAD and the volume of bank loans. As noted, the LEAD is handled to capture the impact of early signals of business cycle fluctuations in an economy on bank lending. The PDP of this indicator is nearly identical to those of GDP in Fig. 6c. Thus, LEAD provides further evidence for the critical role of loan demand and the impact of favorable economic conditions on Turkey's outstanding loans.

Combining the PDPs for GDP and LEAD, one can conclude that an increase in economic activity and the income level in any economy creates an opportunity for successful investment projects and encourages firms to increase their investment projects by demanding more bank loans (Kashyap et al. 1993). Therefore, our findings on economic activity indicators are also consistent with those in many existing studies for various countries (Hoffman 2001; Calza et al. 2003; Imran and Nishat 2013; Sarath and Van Pham 2015; Kapounek et al. 2017) and for Turkey (Akinci et al. 2012; Alper et al. 2012; Macit 2012). It should also be noted that the correlation matrix directed us to handle the fourth lag of GDP and the current values of LEAD in our specification.

Figure 6d, e, f illustrate the relationship between interest rates and outstanding bank loans in Turkey. Among them, Fig. 6d displays the PDP for the interbank ON interest rate to demonstrate the monetary policy's impact on bank's lending behavior. As noted, the underlying mechanism, in which monetary policy changes exert their effects on real economic activity via bank loans, is labeled as the bank lending channel (Çavuşoğlu 2002). The PDP for the ON interest rate in the figure above demonstrates the negative relationship between the monetary policy indicator and bank loans. This

means that banks offer fewer loans during the monetary contraction period. This finding is likely to provide a signal for an active bank lending channel in Turkey, since monetary authorities could affect bank loans via policy rate decisions.

Our findings are consistent with those presented in the studies by Aktaş and Taş (2007) and Alper et al. (2012), favoring the bank lending channel, and are partially in line with Akinci et al. (2012) who proposes a signal for active bank lending channel in the Turkish economy. Further, these results are consistent with the findings in other studies that exhibit a significant negative relationship between policy rates and bank lending (Hoffman 2001; Brissimis et al. 2014; Cantero-Saiz et al. 2014; Brinkmeyer 2014). Unlike existing evidence, our findings provide a nonlinear relationship between the policy rate and bank lending. Furthermore, it can be concluded that policy rate changes have a more significant impact on bank loans.

The PDP for GVT2Y demonstrates that the outstanding negative relationship between the long-term rate measure and bank loans results in an equilibrium after GVT2Y exceeds 20%. The following rise in the long-term market-rate creates a new equilibrium for the current association. These findings support those by Sarath and Van Pham (2015).

In addition, Fig. 6f provides evidence supporting the negative relationship between the 9-month Treasury bill rate and bank lending. Thus, it seems that the rise in short-term market rates negatively affects bank lending behavior. Our finding on the outstanding relationship between short-term market rate and bank lending is consistent with that by Panagopoulos and Spilliotis (1998).

These results are also in line with the illustrated PDPs for LENDINGR, ON, and GVT2Y. The consistency in the relationship between bank lending behavior and various interest rates could be attributed to the central bank policy rate's pass-through to the various market interest rates. Any change in the policy rate seems to be transmitted into the bank lending rate and other market rates accurately, which ensures consistency in the relationship between bank loans and interest rates.

Figure 6g illustrates the PDP for the REER and confirms the negative relationship between REER and bank loans. The rise in the REER demonstrates the appreciation of the domestic currency in our specification. Therefore, one can infer that the appreciation of domestic currency decreases bank lending gradually. There are two possible explanations for this result. First, the domestic currency's appreciation contributes by some valuation effect on the nominal value of bank loans in terms of domestic currency (Imran and Nishat 2013). Another possible explanation for this could be that domestic currency appreciation raises the demand for imports, and economic agents need more bank credit to finance their imports (Shijaku and Kalluci 2013). These results are consistent with those proposed by Ladime et al. (2013) and Pham (2015).

Finally, the current study indicates the relationship between global factors and outstanding bank loans in Turkey. As a small open economy, Turkey might be affected by global economic conditions. Thus, Fig. 7 illustrates the PDPs for ECB main refinancing rate, FFR, and oil prices.

Among the global factors, the ECB main refinancing rate matter most for outstanding bank loans. The PDP in Fig. 7a indicates the nonlinear relationship between ECB and bank lending. The 2% seems to be a threshold level in which the secularly

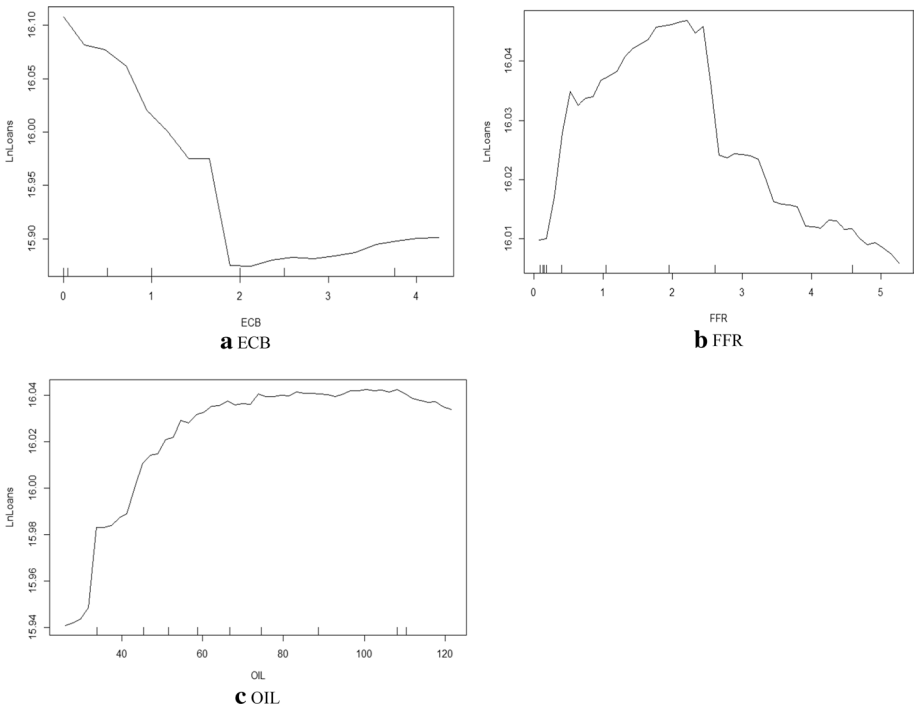

Fig. 7 Partial dependence plots for global factors

decreasing relationship changes to a gradually increasing process. The first part of an outstanding relationship supports the idea that the rise in the ECB refinancing rate increases the financing costs of banks in Turkey. Thus, they might cut their lending. The negative relationship between ECB and bank loans is consistent with Ivanovic (2016), who supported the crucial role of banks' borrowing cost on bank lending. However, the gradual rise in bank loans, after the ECB refinancing rate increases above the 2% level, might be an indicator of the favorable demand conditions in Euro Area countries. The favorable economic conditions in the Euro Area countries, which are Turkey's biggest trade partners, increase their export demand from Turkish firms. In turn, firms might increase their demand for bank loans to enlarge their production and finance their operations.

The PDP for FFR provides different results. Although Fig. 7b demonstrates a nonlinear relationship between FFR and the volume of bank loans, the rise in the FFR, up to slightly above the 2% level, increases Turkey's outstanding bank loans. The FFR, as an indicator of global liquidity, implies a somewhat nonlinear impact on bank lending in Turkey. It seems that the global liquidity drain associated with the rise in FFR causes bank loans to decline in Turkey after the FFR exceeds the 2% level. Further, the apparent discrepancy on the bank loans in Turkey, between ECB and FFR, might be due to the variation in the transmission of the policy rate to the market rates and their ability to affect real economic activity in the Euro Area and the United States.

Finally, Fig. 7c illustrates the relationship between oil prices and bank lending. As noted, oil prices might be an indicator of global economic conditions. They could affect

bank loans in Turkey by changing firms' production costs, altering cash flows, and influencing investment decisions.

The PDP for OIL suggests a steadily increasing relationship between oil prices and bank loans in Turkey. Thus, it is safe to conclude that banks benefit from rising oil prices and increase their lending. These results are likely to be related to demand factors. Since oil price hikes might increase firms' production costs and reduce their cash flows (Kocaarslan and Soytas 2019), firms increase their demand for bank credit.

An alternative explanation for the relationship between oil prices and nominal bank loans in Turkey might be the pass-through of oil prices to the CPI. To illustrate, the rise in oil prices causes the domestic price level to increase, which transmits into the nominal valuation of bank loans. The existing literature provides evidence on the pass-through of oil prices to CPI in Turkey (Çatik and Önder 2011; Çatık and Karaçuka 2012; Akçelik and Öğünç 2016).

Overall, the current study's model specification further provides a comprehensive picture of the drivers of bank lending in Turkey, benefits from the distinct advantages of having machine learning techniques, and provides favorable evidence for some nonlinearities on the outstanding relationship between several indicators and bank lending volume.

Briefly, the findings concerning variable importance and the PDPs indicate that the capital ratio, deposit overhang, and the credit risk are favorable measures influencing bank loans in Turkey. Moreover, these findings suggest that CPI, GDP, and the LEAD have greater predictive salience on bank lending behavior than global factors do.

Discussion and conclusion

This study aimed to investigate the relationship between bank loans and a set of bank-specific characteristics, macroeconomic indicators, and global factors in Turkey between 2002Q4 and 2019Q2. The variable specification is based on the studies in the existing literature. To address the critical role of these variables on bank lending behavior, this study benefits from the distinct advantages of machine learning techniques (regression tree, boosting, bootstrap aggregating, random forest, extra-trees, and xgboost predictors).

The study finds that the random forest model has the lowest predicting error, and thus it has the best out of sample fit measure. Regarding the random forest model results on variable importance and the illustrated PDPs, numerous bank-specific characteristics seem to be the more prominent measures that affect the bank lending behavior.

Among the bank-specific characteristics, the capital ratio, reliance on deposit financing, and credit risk are more salient predictors of bank lending. All these bank-specific characteristics imply a nonlinear association with outstanding bank loans. Therefore, bank loans would enhance with an optimal level of bank capital, deposit financing, and credit risk measure. Besides, balance sheet liquidity, asset structure, and profitability have relatively higher predictive power among the bank-specific characteristics and provide somewhat consistent results with existing studies. However, the bank lending rate seems to be relatively silent in predicting bank lending behavior in terms of the rank of the variable importance.

The empirical findings for macroeconomic indicators suggest that the CPI performs better among its counterparts. The CPI demonstrates the valuation effect on nominal bank loans and captures the rise in loan demand during inflationary periods. Further, the PDPs for GDP and the LEAD indicate the bank loan's procyclical behavior and suggest that improving economic conditions would increase the bank lending volume. Furthermore, the relationship between the policy and bank loans favors the bank lending channel in Turkey, despite the ON interbank rate's lower predictive performance. In contrast, the role of global factors in predicting bank lending behavior in Turkey seems less reliable than several bank-specific characteristics and macroeconomic indicators.

In this regard, the current study's findings have some policy implications for bank managers, regulatory authorities, and policymakers. Generally, we may conclude that the higher-ranked indicators are most likely to be more salient policy levers, based on their importance rank, compared to the lower-ranked measures. Thus, the rank of variable importance enables policymakers to be more confident in designing favorable policies. Notably, the nonlinear nature of the link between capital ratio, deposit overhang, and the credit risk force policymakers to design related policies more prudently. The reliable predicting power of bank-specific characteristics on bank lending calls for the banking authorities to consider their statistical reports properly.

The role of bank-specific characteristics on bank lending behavior also highlights the significant role of bank-level managerial decisions. To illustrate, the importance of liquidity and the profitability in predicting outstanding bank loans indicates that focusing bank-level decision support systems on the various risk measures might be beneficial for deposit banks.

Furthermore, the importance of economic activity variables suggests that providing more sound economic growth policies would enhance the volume of bank loans. In addition, the relatively higher importance of GDP and LEAD provide evidence regarding the role of loan demand factors on outstanding bank loans. Further, the relatively critical nature of monetary policy rates requires monetary authorities to be more progressive in designing policies to provide financial stability. The monetary authority might also promote economic activity since these results provide evidence based on an active bank lending channel in Turkey.

For global factors, although their importance is somewhat lower than that of other measures, policymakers should take precautionary measures using available tools when global shocks create conflicting disorders in the banking sector and macroeconomic environment. Finally, these results might also increase policymakers' confidence about negative policies. Since the current study offers the importance ranks of the variables, the results also indicate the policy levers that might be less effective in shaping bank lending behavior.

Since bank loan data for different types of loans (consumer loans, commercial loans, automobile loans) is not available, the sum of all types of bank loans is used for a particular deposit bank. This study is, therefore, limited in that it does not focus on the impact of different factors on various types of bank loans. Further, there is scope for further research, since this study does not consider bank size as a bank-specific characteristic in the model specification. Further research might investigate the potential heterogeneity in the bank lending behavior in banks with different size groupings.

Abbreviations
BAT: Banks Association of Turkey; CBRT: Central Bank of the Republic of Turkey; EDDS: Electronic Data Delivery System; GDP: Gross Domestic Product; GSIFI: Global Systematically Important Financial Institutions; MAPE: Mean absolute percentage error; MSE: Mean square error; NPL: Non-performing loans; ROA: Return on asset; US: United States; PDP: Partial dependence plot; WDI: World development indicators.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40854-021-00237-1.

Additional file 1. Appendix.

Acknowledgements
We are thankful to anonymous referees for their kindly suggestions on this paper.

This manuscript is the revised and improved version of the third chapter of the Ph.D. Dissertation entitled "Three Essays on Applied Macroeconomics" proposed by Onder Ozgur in Ankara Yildirim Beyazit University, Graduate School of Social Sciences.

Authors' contributions
OO: conceptualization, Investigation, Methodology, Data curation, writing-Original draft preparation, Software. ETK: supervision, conceptualization, reviewing and editing. FCO: supervision, conceptualization, reviewing and editing. All authors read and approved the final manuscript.

Funding
We do not receive any financial assistance from any agency.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare no competing interests

Received: 20 April 2020 Accepted: 22 March 2021
Published online: 29 March 2021

References

Akcelik F, Öğünç F (2016) Pass-through of crude oil prices at different stages in Turkey. Central Bank Rev 16(1):41–51. https://doi.org/10.1016/j.cbrev.2016.03.004
Akinci DA, Matousek R, Radic N, Stewart C (2012) Monetary policy and the banking sector in Turkey. Working Paper. London Metropolitan University, London, UK
Aktaş C, Taş BKO (2007) The bank lending channel in Turkey: effect of capital adequacy ratio. J BRSA Bank Financ Marks 1:61–76
Alessi L, Detken C (2018) Identifying excessive credit growth and leverage. J Financ Stab 35:215–225. https://doi.org/10.1016/j.jfs.2017.06.005
Alpaydin E (2014) Introduction to Machine Learning, 3rd edn. The MIT Press
Alper K, Hulagu T, Keles G (2012) An empirical study on liquidity and bank lending. Central Bank Repub Turkey Work Pap 12(4):1–18
Altabrawee H (2017) Performance evaluation of fourteen machine learning algorithms on credit card default classification. J Coll Educ Wasit 1(26):461–474
Athey S (2018) The impact of machine learning on economics. The Economics of Artificial Intelligence: an agenda. University of Chicago Press, pp 507–547
Baoko G, Acheampong IA, Ibrahim M (2017) Determinants of bank credit in Ghana: a bounds-testing cointegration approach. Afr Rev Econ Finance 9(1):33–61
Basuchoudhary A, Bang JT, Sen T (2017) Machine-learning techniques in economics: new tools for predicting econ growth. Springer
Berger AN, Udell GF (1994) Did risk-based capital allocate bank credit and cause a "credit crunch" in the United States? J Money Credit Bank 26(3):585–628. https://doi.org/10.2307/2077994
Bernanke BS (1993) Credit in the macroeconomy. Q Rev-Federal Reserve Bank N Y 18:50–70
Berrospide JM, Rochelle ME (2010) The effects of bank capital on lending: what do we know, and what does it mean? Int J Central Bank 6(4):5–54
Bertay AC, Demirgüç-Kunt A, Huizinga H (2012) Bank ownership and credit over the business cycle: is lending by state banks less procyclical? CEPR Discussion Paper, No. 9034

Blaes, B (2011) Bank-related loan supply factors during the crisis: an analysis based on the German bank lending survey. Deutsche Bundesbank Discussion Paper Series 1, Econ Studies 31
Blundell-Wignall A, Roulet C (2013) Bank lending puzzles. OECD J Financ Mark Trends 1:7–30. https://doi.org/10.1787/fmt-2013-5k40m1nz55wj
Boyd JH, Prescott EC (1986) Financial intermediary-coalitions. J Econ Theory 38(2):211–232
Brainard WC, Tobin J (1963) Financial intermediaries and the effectiveness of monetary controls. Am Econ Rev 53(2):383–400
Brei M, Gambacorta L, Von Peter G (2013) Rescue packages and bank lending. J Bank Finance 37(2):490–505. https://doi.org/10.1016/j.jbankfin.2012.09.010
Breiman L (1996) Bagging predictors. Mach Learn 24(2):123–140. https://doi.org/10.1007/BF00058655
Breiman L (2001) Random forests. Mach Learn 45(1):5–32
Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. CRC Press, Boca Raton
Brinkmeyer H (2014) Drivers of bank lending: new evidence from the crisis. Springer
Brissimis S, Delis M, Iosifidi M (2014) Bank mark power and monetary policy transmission. Int J Central Bank 10:173–213. https://doi.org/10.2139/ssrn.2858468
Calza A, Gartner C, Sousa J (2003) Modelling the demand for loans to the private sector in the euro area. Appl Econ 35(1):107–117. https://doi.org/10.1080/00036840210161837
Cantero-Saiz M, Sanfilippo-Azofra S, Torre-Olmo B, López-Gutiérrez C (2014) Sovereign risk and the bank lending channel in Europe. J Int Money Finance 47:1–20
Carlson M, Shan H, Warusawitharana M (2013) Capital ratios and bank lending: a matched bank approach. J Financ Intermed 22(4):663–687. https://doi.org/10.1016/j.jfi.2013.06.003
Casu B, Girardone C, Molyneux P (2006) Introduction to Banking. Vol. 10. Pearson Education.
Chen T, Guestrin C (2016) Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp 785–794
Clark GL (2017) Financial intermediation, infrastructure investment and regional growth. Area Dev Policy 2(3):217–236. https://doi.org/10.1080/23792949.2017.1345641
Cucinelli D (2015) The impact of non-performing loans on bank lending behavior: evidence from the Italian Bank sector. Eurasian J Bus Econ 8(16):59–71. https://doi.org/10.17015/ejbe.2015.016.04
Çatik AN, Karaçuka M (2012) Oil pass-through to domestic prices in Turkey: does the change in inflation regime matter? Econ Res Ekonomska Istraživanja 25(2):277–296. https://doi.org/10.1080/1331677X.2012.11517508
Çatik AN, Önder AÖ (2011) Inflationary effects of oil prices in Turkey: a regime-switching approach. Emerg Marks Finance Trade 47(5):125–140. https://doi.org/10.2753/REE1540-496X470506
Çavuşoğlu AT (2002) Credit transmission mechanism in Turkey: an empirical investigation. Middle East Technical University. ERC Work Pap Econ 2(3):1–30
Demiralp S, Eisenschmidt J, Vlassopoulos T (2017) Negative interest rates, excess liquidity and bank business models: Banks' reaction to unconventional monetary policy in the euro area. Excess Liquidity and Bank Business Models: Banks' Reaction to Unconventional Monetary Policy in the Euro Area. SSRN Working Paper (2017)
Diamond DW (1984) Financial intermediation and delegated monitoring. Rev Econ Stud 51(3):393–414
EDDS (2020) Money and Banking Statistics. https://evds2.tcmb.gov.tr/index.php?/evds/serieMarket. Accessed 5 Mar 2020.
Égert B, Backé P, Zumer T (2007) Private-sector credit in Central and Eastern Europe: New (over) shooting stars?. Comp Econ Stud (49):201–231
Fornaro P, Luomaranta H (2020) Nowcasting Finnish real economic activity: a machine learning approach. Empir Econ 58(1):55–71. https://doi.org/10.1007/s00181-019-01809-y
Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 55:119–139. https://doi.org/10.1007/3-540-59119-2_166
Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29(5):1189–1232
Gambacorta L, Marques-Ibanez D (2011) The bank lending channel: lessons from the crisis. Econ Policy 26(66):135–182. https://doi.org/10.1111/j.1468-0327.2011.00261.x
Gambacorta L, Rossi C (2010) Modelling bank lending in the euro area: a nonlinear approach. Appl Financ Econ 20(14):1099–1112. https://doi.org/10.1080/09603101003781430
Gertler M, Gilchrist S (1993) The role of credit market imperfections in the monetary transmission mechanism: arguments and evidence. Scand J Econ 95(1):43–64. https://doi.org/10.2307/3440134
Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. Mach Learn 63(1):3–42. https://doi.org/10.1007/s10994-006-6226-1
Greenbaum SI, Thakor AV, Boot AW (2019) Contemporary financial intermediation, 4th edn. Academic Press
Greenwell BM (2017) pdp: An R package for constructing partial dependence plots. R J 9(1):421
Hastie T, Tibshirani R, Friedman J (2017) The elements of statistical learning: data mining, inference, and prediction. Springer
Hofmann B (2001) The determinants of private sector credit in industrialised countries: do property prices matter? BIS working papers, vol 108, pp 1–39
Imran K, Nishat M (2013) Determinants of bank credit in Pakistan: a supply side approach. Econ Model 35:384–390. https://doi.org/10.1016/j.econmod.2013.07.022
Ivanović M (2016) Determinants of credit growth: the case of Montenegro. J Central Bank Theory Pract 5(2):101–118. https://doi.org/10.1515/jcbtp-2016-0013
James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning, vol 112. Springer
Jiménez G, Ongena S, Peydró JL, Saurina J (2012) Credit supply and monetary policy: identifying the bank balance-sheet channel with loan applications. Am Econ Rev 102(5):2301–2326. https://doi.org/10.1257/aer.102.5.2301
Kapounek S, Kučerová Z, Fidrmuc J (2017) Lending conditions in EU: the role of credit demand and supply. Econ Model 67:285–293. https://doi.org/10.1016/j.econmod.2017.01.003
Kara H (2016) A brief assessment of Turkey's macroprudential policy approach: 2011–2015. Central Bank Rev 16(3):85–92

Kashyap AK, Stein JC (1995) The impact of monetary policy on bank balance sheets, Carnegie-Rochester Conference Series on Public Policy, vol 42
Kashyap A, Stein J, Wilcox D (1993) The monetary transmission mechanism: evidence from the composition of external finance. Am Econ Rev 83(1):78–98
Kim D, Sohn W (2017) The effect of bank capital on lending: does liquidity matter? J Bank Finance 77:95–107. https://doi.org/10.1016/j.jbankfin.2017.01.011
Kocaarslan B, Soytas U (2019) Asymmetric pass-through between oil prices and the stock prices of clean energy firms: new evidence from a nonlinear analysis. Energy Rep 5:117–125. https://doi.org/10.1016/j.egyr.2019.01.002
Košak M, Li S, Lončarski I, Marinč M (2015) Quality of bank capital and bank lending behavior during the global financial crisis. Int Rev Financ Anal 37:168–183. https://doi.org/10.1016/j.irfa.2014.11.008
Ladime J, Sarpong-Kumankoma E, Osei KA (2013) Determinants of bank lending behavior in Ghana. J Econ Sustain Dev 4(17):42–47
Leo M, Sharma S, Maddulety K (2019) Machine learning in banking risk management: a literature review. Risks 7(29):1–22
Levine R (2005) Finance and growth: theory and evidence. Handb Econ Growth 1:865–934. https://doi.org/10.1016/S1574-0684(05)01012-9
Macit F (2012) Who responds more to monetary policy, conventional banks or participation banks? J Econ Finance Admin Sci 17(33):10–14. https://doi.org/10.1016/S2077-1886(12)70003-3
Malede M (2014) Determinants of commercial banks lending: evidence from Ethiopian commercial banks. Eur J Bus Manag 6(20):109–117
Mullainathan S, Spiess J (2017) Machine learning: an applied econometric approach. J of Econ Perspectives 31(2):87–106. https://doi.org/10.1257/jep.31.2.87
Naceur SB, Marton K, Roulet C (2018) Basel III and bank-lending: evidence from the United States and Europe. J Financ Stab 39:1–27. https://doi.org/10.1016/j.jfs.2018.08.002
Orlova EV (2020) Decision-making techniques for credit resource management using machine learning and optimization. Information 11(3):144
Panagopoulos Y, Spiliotis A (1998) The determinants of commercial banks' lending behavior: some evidence for Greece. J Post Keynes Econ 20(4):649–672. https://doi.org/10.1080/01603477.1998.11490172
Petropoulos A, Siakoulis V, Stavroulakis E, Klamargias A (2019) A robust machine learning approach for credit risk analysis of large loan level datasets using deep learning and extreme gradient boosting. IFC Bulletins chapters. In: Bank for International Settlements (ed) Are post-crisis statistical initiatives completed?, p 49
Pham THH (2015) The Determinants of Bank Lending. HAL working paper, 01158241l, Universidad de Nantes. https://hal.archives-ouvertes.fr/hal-01158241/document. Accessed 2 Jan 2020
Potjagailo G (2017) Spillover effects from Euro area monetary policy across Europe: a factor-augmented VAR approach. J Int Money Finance 72:127–147. https://doi.org/10.1016/j.jimonfin.2017.01.003
Rabab'ah M (2015) Factors affecting the bank credit: an empirical study on the Jordanian commercial banks. Int J Econs Finance 7(5):166–178
Ramey V (1993) How important is the credit channel in the transmission of monetary policy? Carnegie-Rochester Conference Series on Public Policy, North Holland, no 39, pp 1–45. North-Holland
Sarath D, Van Pham D (2015) The determinants of Vietnamese banks' lending behavior. J Econ Stud 42(5):861–877. https://doi.org/10.1108/JES-08-2014-0140
Seven Ü, Yetkiner H (2016) Financial intermediation and economic growth: does income matter? Econ Syst 40(1):39–58. https://doi.org/10.1016/j.ecosys.2015.09.004
Shahzad F, Fareed Z, Zulfiqar B, Habiba U, Ikram M (2019) Does abnormal lending behavior increase bank riskiness? Evid Turkey Financ Innov 5(1):37
Shen F, Zhao X, Kou G (2020) Three-stage reject inference learning framework for credit scoring using unsupervised transfer learning and three-way decision theory. Decis Support Syst 137:113366
Shijaku G, Kalluci I (2013) Determinants of bank credit to the private sector: the case of Albania. Bank of Albania Working Paper, No. 09(48), pp 1–36
Tang J, Zheng L, Han C, Liu F, Cai J (2020) Traffic incident clearance time prediction and influencing factor analysis using extreme gradient boosting model. J Adv Transp. https://doi.org/10.1155/2020/6401082
Tuffery S (2011) Data mining and statistics for decision making. Wiley series in computational statistics
Us V (2017) Dynamics of non-performing loans in the Turkish banking sector by an ownership breakdown: the impact of the global crisis. Financ Res Lett 20:109–117. https://doi.org/10.1016/j.frl.2016.09.016
Varian H (2014) Big data: new tricks for econometrics. J Econ Perspect 28(2):3–28. https://doi.org/10.1257/jep.28.2.3
Varlik S, Berument MH (2017) Multiple policy interest rates and economic performance in a multiple monetary-policy-tool environment. Int Rev Econ Finance 52:107–126. https://doi.org/10.1016/j.iref.2017.10.004
Wang M, Yu J, Ji Z (2018) Credit fraud risk detection based on XGBoost-LR hybrid model.
Wang H, Kou G, Peng Y (2020) Multi-class misclassification cost matrix for credit ratings in peer-to-peer lending. J Oper Res Soc 1–12. https://doi.org/10.1080/01605682.2019.1705193
Zhu L (2019) Predictive modelling for loan defaults, Doctoral dissertation, UCLA

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.