Global Finance Journal 63 (2024) 101050

Contents lists available at ScienceDirect

Global Finance Journal

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

Estimating probability of default via delinquencies? Evidence from European P2P lending market

Asror Nigmonov[a,*], Syed Shams[b], Povilas Urbonas[c]

[a] University of New South Wales, Sydney, Australia
[b] University of Southern Queensland, Australia
[c] Startup Wise Guys, Vilnius, Lithuania

A R T I C L E I N F O

JEL classification:
E31
E43
G29
G14
O16

Keywords:
Peer-to-peer lending
FinTech
Default
Marketplace lending
Panel data: LASSO method

A B S T R A C T

The unprecedented growth of the financial sector's digital transformation opens wide areas to the scaling up of finance in innovative and knowledge-based projects. Improving risk management takes centre stage in the acceleration of this process. This study uses loan-book data from the peer-to-peer (P2P) lending market to empirically investigate the determinants of default risk. Using the loan-book database covering the period from 2014 to 2020, we examine multiple factors related to the default risk of loans issued by P2P lending platforms. The results indicate that a higher interest rate and higher stock market returns increase the probability of default in the P2P lending market. Results are robust to additional tests based on endogeneity correction, the LASSO method and sampling bias. The severity of the impact of market returns and interest rates is found to be significantly different based on the levels of financial technology (FinTech) adoption and banking sector distress. Increases in the market interest rate are found to boost the sensitivity of P2P loan defaults to stock market volatility. This study contributes to existing literature on risk management models with its consideration of country-specific factors, paving the way to future best practices in the market.

1. Introduction

The rapid advancement of digital technologies has fundamentally changed the competitiveness of financial sectors worldwide, accelerating the rise of digital payments, mobile banking blockchain contracts, digital currency and the interconnected digital ecosystems of financial sectors. The capital available to more efficient firms could be reduced by the accelerated growth of financial technology (FinTech) (Xie & Zhu, 2022). Furthermore, these developments in financial technology could shift investors' preferences towards non-traditional markets which are less liquid and transparent (Sindreu, 2020). This may allow a new generation of non-banks to come of age, transforming shadow banking (banking by non-banks) within the broader development of digital financial industries.

Peer-to-peer (P2P) lending is one form of alternative lending practices transformed into a new financing channel for underserved communities and entrepreneurial start-ups (Oren, 2013; Zhang et al., 2016). This form of lending is part of a broader financial market where market and default risks are tied together, with P2P lending also considered to be a subcategory of FinTech lending in which borrowers and lenders directly interact via the online platform. Throughout this paper, we use P2P lending in relation to discussion of our findings. However, we compare our study's findings to broader FinTech lending where relevant.

* Corresponding author.
E-mail addresses: a.nigmonov@unsw.edu.au (A. Nigmonov), syed.shams@usq.edu.au (S. Shams), povilas@startupwiseguys.com (P. Urbonas).

https://doi.org/10.1016/j.gfj.2024.101050
Received 25 December 2023; Received in revised form 2 October 2024; Accepted 10 October 2024
Available online 11 October 2024
1044-0283/

Several studies (Dyreng et al., 2012; Stiglitz & Weiss, 1981, 1992) have provided evidence that default in the traditional finance model depends on factors such as the market's country-specific demographics and economic conditions. However, determining the factors behind borrowers' probability of default in the FinTech lending market is important due to the critical trade-off between profit margin and default risk faced by Fintech lenders (Lin et al., 2023). These platforms also need to accurately estimate the probability of loan default to effectively match the demand for and supply of funds, with this not being the case in traditional banking (Nigmonov et al., 2022).

Accordingly, the current study aims to extend this strand of the literature by examining the specific determinants of default risk on P2P platforms in relation to the broader financial market. Specifically, this study aims to answer the following central question: *How is the probability of default determined* via *stock market returns and interest rates?*

This research question is significant as it explores the interconnectedness of financial markets and their impact on default probability. The extant traditional finance literature has indicated that higher interest rates tend to increase the default risk by limiting borrowers' ability to meet their debt obligations (Bester, 1985; Stiglitz & Weiss, 1981, 1992). Nigmonov et al. (2022); Serrano-Cinca et al. (2015); and Wei and Lin (2016) highlighted that higher interest rates increase the probability of loan default in P2P lending. Interest rates were also found to directly affect the cost of borrowing and, thus, may impact the likelihood of delinquencies and defaults (Bergmann, 2020; Fuster & Willen, 2017). By investigating these relationships, the current study provides valuable insights into risk management strategies in the European P2P lending market and in the broader financial sector.

This study utilises the loan-book of Mintos (Latvia), one of the leading P2P lending platforms in Continental Europe. This platform operates across the multiple countries of the European Union (EU), making it ideal for exploring the default risk incurred by P2P lending platforms in Europe. Europe is considered as a significant target research jurisdiction for several fundamental reasons. Firstly, the European P2P lending market, being a relatively new and rapidly growing sector, provides a rich data set for studying default probability. Specifically, the Nordic–Baltic region[1] performs exceptionally well with significant increases in FinTech usage, growth in financial capability and growth in FinTech scores (Nourallah et al., 2024). The current Mintos database covers 10 countries of Europe; hence, it has the largest geographical scope known to authors. Secondly, the study of P2P lending in Europe is necessitated by P2P lending's unique characteristics and the transformative impact it has on the financial landscape. Specifically, existing studies have highlighted the greater reliance on non-traditional factors (e.g., soft information derived from the description text of the loan application) by European P2P lenders (Berg et al., 2022; Dorfleitner et al., 2016).

Thirdly, P2P lending platforms in Europe have been serving high-quality and creditworthy small businesses that already have access to bank credit (Eça et al., 2021). In the context of Europe, the ability of financial technologies to set up and access capital in combination with their growth has been a key focus. These facts highlight the vibrant FinTech ecosystem in Europe, driven by a combination of progressive regulations, sustainable product offerings, technological advancements, and a strong appetite for digital transformation (Cambridge Centre for Alternative Finance (CCAF), 2024). Fourthly, the importance of the European FinTech market is underscored by its rapid growth and potential for further expansion. However, the maturity and performance of FinTech ecosystems by country are widely divergent, with substantial gaps between European countries (McKinsey & Company, 2022). Finally, FinTech in Europe has limited regulation compared to other regions, such as the United States (US) and Asia. Europe stands out as the region for which the highest percentage (8 %) of FinTech company representatives state that no specific regulation exists nor is it needed, according to the survey conducted by Cambridge Centre for Alternative Finance (Cambridge Centre for Alternative Finance (CCAF), 2024). Therefore, Europe as a case provides a unique institutional perspective for generalising and scaling the findings of this research globally.

Our study's results, consistent with the traditional banking literature's theoretical predictions (Bester, 1985; Stiglitz & Weiss, 1981, 1992), show that higher interest rates and higher stock market returns lead to higher default rates. Our study's main findings remain robust for endogeneity correction based on instrumental variables, the least absolute shrinkage and selection operator (LASSO) method and sampling bias. Our study also indicates the transmission of credit risk from traditional financial markets to FinTech lending markets. The magnitudes of these spillovers are found to be significantly different, depending on the broader interest rate environment and distress levels in the banking sector as well as FinTech adoption and borrower ratings. Finally, our findings indicate the impact of market returns and interest rates on the survivability of loans with respect to the loan closure date, with loans highly sensitive to market forces during the first 100 days from their issue date.

This study's findings contribute to the existing literature in several ways. Firstly, the study conducts estimations in a multi-country setting which is not attempted by prior studies, such as Nigmonov et al. (2022); Serrano-Cinca et al. (2015); and Wei and Lin (2016). Secondly, this study's focus allows a better understanding to be gained of risk transmission between traditional and FinTech lending on the European continent. Thirdly, our study's findings improve risk management models by better estimating the country risk component. Finally, given the global expansion of the P2P lending industry and the growing interest in this topic, the study has important implications for practitioners and researchers by explicitly linking interest rates and stock market returns with the default risk of P2P borrowers.

The rest of this paper proceeds as follows: in Section 2, the literature review is discussed and the study's hypotheses are developed. Section 3 highlights the regression method used in this study. Section 4 discusses the empirical results, while Section 5 concludes the paper with our study's implications, the generalisability of the study's findings and directions for further research.

[1] Including Estonia, Finland, Latvia and Sweden in the analysis in this study.

2. Literature review and hypotheses development

2.1. Review of existing literature related to default risk

The development of default risk models largely stems from theoretical foundations evaluating the need to estimate defaults among business entities. Evaluating the probability of default (PD) is crucial when assessing credit exposure and potential losses faced by firms. Corporate securities carry the risk of eventual default during the contract term. Merton's (1974) model plays a fundamental role in estimating default risk for firms by establishing a connection between default risk and a firm's capital structure. On the other hand, Altman (1968) pioneered the use of financial ratios and discriminant analysis for predicting corporate bankruptcy (Altman's *Z*-Score model). His work was further extended by Ohlson (1980) who introduced a probabilistic model with logit analysis incorporating financial ratios for bankruptcy prediction.

The later study by Jones (2017) conducted a high-dimensional analysis of corporate bankruptcy prediction, highlighting the importance of considering a broad set of variables, including corporate governance. Liang et al. (2016) conducted a comprehensive study on bankruptcy prediction, incorporating financial ratios and corporate governance indicators. Their findings underscored the importance of effective corporate governance in mitigating bankruptcy risk. Similarly, Ooghe and De Prijcker (2008) developed a typology of company bankruptcy, emphasising the need to understand the underlying processes and causes of failure.

Generally, researchers have categorised corporate default prediction models into three generations: discriminant analyses, binary response models and hazard models (Kim et al., 2020). The modelling framework of the current study originates from the second generation in baseline analysis that primarily focuses on binary response models. This group of models do not assume specific distributions for predictor variables and allow the significance of individual independent variables to be tested. The third-generation models use a duration analysis or survival analysis to demonstrate superior prediction performance (Chava & Jarrow, 2004; Shumway, 2001). The studies of Duan et al. (2012); Nam et al. (2008); and Tian et al. (2015) used third-generation models with industry effects, market variables and macroeconomic dependencies, largely outperforming conventional models pioneered by Altman (1968). Following this evidence from existing studies, we support our baseline results with survival analysis in Section 4.5.

Beyond traditional models, recent studies have explored machine learning techniques for default prediction. These approaches leverage large data sets, feature engineering and algorithmic models to enhance the probability of default (PD) estimation. Studies by Chen (2011) and Shin et al. (2005) identified the high level of accuracy and better performance of machine learning methods for short- and long-term default. Jabeur et al. (2021) explored the use of the CatBoost model, with artificial intelligence (AI) techniques for corporate failure prediction based on gradient boosting decision trees. The studies of Babaei et al. (2023); Tyagi (2022); and Xia et al. (2021) employed the latest applications of advanced machine learning methods for default estimation. Their research focused on enhancing default prediction accuracy, by leveraging machine learning algorithms and feature engineering.

Ogundimu (2022) highlighted that supervised machine learning models are particularly suitable for variable selection, with the adaptive LASSO method demonstrating superior combined effects on sensitivity and specificity, compared to *p*-value-based methods which are threshold dependent. LASSO models have been found applicable for models with high dimensionality and can handle many potential covariates while yielding a parsimonious specification (González-Coya & Perron, 2024; Mei & Shi, 2024). Ogundimu (2022) suggested that LASSO methods should be preferred for achieving optimal predictions in the context of FinTech loan defaults. Jiang et al. (2021) indicated that LASSO models successfully predicted default and outperformed other competitive models for both in-sample and out-of-sample tests. Based on this background discussion, the current study uses the supervised machine learning LASSO method to improve the default estimation presented in the baseline regression.

2.2. Default risk in P2P lending

The rise of the P2P lending market is often attributed to traditional bank lending regulation after the Global Financial Crisis (GFC). Even though the P2P lending process does not use the traditional banking concept, P2P lending platforms mediate between borrowers and lenders by charging a service fee. Thus, the primary purpose of P2P lending platforms is to match the demand for and supply of funds by correctly estimating the probability of loan default (Wei & Lin, 2016). At the same time, the growing problem of P2P lending platforms is the credit default risk due to information asymmetry (Emekter et al., 2015).

In P2P lending markets, lenders are not informed about the risks, with borrowers' credit conditions mainly dependent on the extent of borrowers' knowledge about themselves. Once the transaction has taken place, lenders may not be able to observe borrowers' actions, project return or duration, or may not be able to force borrowers to repay loans due to ex-post asymmetric information problems (Jaffe & Stiglitz, 1990). As borrowers only negotiate with lenders in an online environment and do not meet face to face, it is difficult for lenders to know borrowers' credit information and debt status (Weiss et al., 2010).

Following the findings of Akerlof (1970) and Greenwald and Stiglitz (1987), these types of situations may lead to adverse selection and moral hazards. Drawing on insights from asymmetric information theory, several models have included possible scenarios where macroeconomic variables affect the net worth of potential borrowers (Bebczuk, 2003). When the asymmetric information problem arises, it causes the 'financial accelerator' effect, with the relationship possibly complicated and highly intractable (Bernanke et al., 1998).

In its analysis of the risk of default in P2P lending markets, the current study uses a range of control variables to represent various factors that affect default risk. On the one hand, financial accelerator theory posits that borrowing becomes more difficult and expensive during a recession, owing to increased external finance premiums (Bernanke & Gertler, 1989; Kiyotaki & Moore, 1997). On the other hand, life-cycle consumption theory highlights that when income is low, people borrow more against an uncertain future

(Lawrence, 1995). Enhanced levels of borrowing, in turn, increase consumption rates and borrowers' debts (Lawrence, 1995). Félix et al. (2013) and Gompers and Lerner (1998) empirically tested these theories in alternative investment markets by proposing that stock market development, as well as growth in gross domestic product (GDP), inflation and unemployment have an impact on the development of venture capital investments. The empirical models used in the current study are drawn from these concepts and use macroeconomic variables (e.g., GDP growth, inflation) as controls for determining loan defaults.

The characteristics of traditional financial markets, such as banks, are also likely to play a part in shaping the level of risk of FinTech loans. More concentrated banking sectors may lead to high lending costs and inefficient processes for borrowers. Thus, in banking sectors with high lending rates and margins, FinTech credit can act as a substitute (Hodula, 2022). In this regard, bank loans may be preferred over FinTech loans due to the endogenous advantage of banks; however, based on financial intermediation theory, FinTech loans may have a cost advantage (Merton & Thakor, 2019; Thakor, 2020).

On the other hand, due to negative credit supply shocks, a substantial portion of borrowers who fall outside the traditional credit supply chains may seek alternative lending markets (Cornett et al., 2011). Ma et al. (2023) discovered that, in concentrated banking markets, borrowers in P2P lending are generally more creditworthy. This indicates the migration of borrowers from banks to lending marketplaces. In this case, low-quality borrowers are forced to seek FinTech credit after being denied access to bank credit, leading FinTech lenders to go 'bottom fishing' (de Roure et al., 2022). Moreover, regulation has been proven to be a mixed blessing for banks, with regulatory shocks not only creating a competitive disadvantage, but also being a barrier to entry (Stulz, 2019).

2.3. Hypotheses development

The existing literature has emphasised the importance of online P2P lending platforms as a convenient alternative means of financing for individual investors (Tang, 2019; Thakor, 2020), as acquiring small loans from traditional financial markets could be difficult or costly. Boubaker et al. (2015) also indicated that small firms, compared to large firms, tend to have greater price reversals in the immediate aftermath of unexpected events, such as political tensions. Thus, a boom in the stock market can increase the number of risk-taking borrowers in the P2P lending market (Yoon et al., 2019). Theoretical studies (Collin-Dufresne & Goldstein, 2001; Duffie et al., 2007) as well as empirical studies (Ericsson et al., 2009; Giesecke et al., 2011; Norden & Weber, 2009) on bank credit markets have indicated that increased volatility in stock market returns raises the probability of credit default.

Specifically, the trade-off between relationship loans and transaction loans has been well documented, with banks preferred by those seeking relationship loans (Boot & Thakor, 2000; Degryse & Ongena, 2007). Accordingly, FinTech lenders involved in relationship lending will be less competitive than banks (Balyuk & Davydenko, 2019), bringing P2P and other FinTech loans closer to transaction loans and, thus, to the capital market end of the spectrum (Thakor, 2020).

While existing studies have indicated the impact of crucial factors, such as borrower-specific characteristics, the macro environment and market competition, other issues may potentially influence the platform default risk. As highlighted by Yoon et al. (2019), one such factor could be the systemic risk inherent in the loan pool. FinTech lending platforms may allow low credit borrowers to borrow money under higher lending rates. Systemic risk can increase if these low credit borrowers fail to repay their loans, spreading the contagion to other platforms and leading to collective failure. Previous studies in the banking literature have explored this contagion phenomenon (Akhtaruzzaman et al., 2021; Allen & Gale, 2000; Gai & Kapadia, 2010), demonstrating that the spread of contagion depends on the pattern of interconnectedness between banks. With a priori indistinguishable shocks having vastly different consequences.

The current study, in exploring this potential systemic risk in the P2P lending market, underscores its unique contribution to the existing body of knowledge. Therefore, this study proposes the following hypothesis on the relationship between the stock market and the P2P lending market:

H1. A high rate of stock market returns increases the probability of default in peer-to-peer (P2P) lending markets.

The theoretical evidence from studies on traditional financial markets has demonstrated the relationship between the interest rate and the probability of default. Firstly, in a high interest rate environment, lenders charge a higher premium, creating a heavier burden on borrowers to service the loan. Secondly, debt servicing becomes more expensive when real interest rates rise and the real value of borrowers' debt increases. Consequently, higher interest rates tend to limit borrowers' ability to meet their debt obligations and, thus, increase the default risk (Bester, 1985; Stiglitz & Weiss, 1981, 1992). Based on the traditional finance literature, this study expects higher interest rates to increase the default risk on P2P lending platforms:

H2. A higher interest rate increases the probability of default in peer-to-peer (P2P) lending markets.

3. Model description

As mentioned in Section 2, the current study relies on binary response models, classified as second-generation models by Kim et al. (2020), for default estimation. This type of model has been widely used in the estimation of defaults and default premium, as pioneered by Aretz et al. (2018) and Campbell et al. (2008). These models do not require assumptions about the probability of default or the distributions of predictor variables and can test the significance of individual independent variables.

The empirical technique used in the current study is logit regression analysis which estimates the dependent variable with a binary value. Logit regression is widely used in credit risk applications (both in practice and research); therefore, it is well understood and comparatively easy to communicate. Thus, replicating the later applications of logit models to estimate defaults, in studies such as those by Bonfim (2009) and Kukuk and Rönnberg (2013), our study's model is developed to estimate default risk as per Eq. (1).

The key dependent variable is the binary variable representing the status of each loan. Logit regression estimates the determinants of the probability of default via the variable *BADDEBT*:

$$BADDEBT_{i,t} = \alpha + \beta_1 X_{i,t-6} + \beta_2 POPESTIMATE_{i,t} + \beta_3 LOANVOL_{i,t-12} + \beta_4 AVEAMNT_{i,t} + \beta_5 AVEDUR_{i,t} + \beta_6 AVERATING_{i,t} + \beta_7 EARNINGS_{i,t} + \beta_8 GDP_{i,t} + \beta_9 AVERATING_{i,t} + \beta_{10} ESI_{i,t} + \beta_{11} INFLATION_{i,t} + \beta_{12} NPL_{i,t} + \beta_{13} INTUSE_{i,t} + e_{it} \tag{1}$$

where $BADDEBT_{i,t}$ is the binary variable representing the loan status with a value of 1 if the loan is classified as a bad loan,[2] and 0 otherwise. In Eq. (1), $X_{i,t-6}$ is one of the two independent variables of interest: (1) monthly average value of daily squared stock market returns at time *t*–*6* in country *i* ($MARKETRET_{i,t-6}$) and (2) interest rate set for issued loans by Mintos for borrower *i* at time *t* ($INTRATE_{i,t-6}$); while e_{it} is the error term.

The existing literature has shown that changes in macroeconomic determinants affect loan quality with different time lags (Ali & Daly, 2010; Norden & Wagner, 2008). Specifically, real GDP, nominal interest rates and inflation have an impact on loan quality with a lag of at least two quarters (Bocola et al., 2019; Chen & Wu, 2014). From a theoretical standpoint, based on business cycles, the ability of households to repay loans is affected by macroeconomic changes after their income is impacted, typically with a lag (Rubaszek & Serwa, 2014). Macroeconomic indicators, such as GDP growth, are themselves lagging business cycle indicators.

Following Bocola et al. (2019) and Chen and Wu (2014), we use lagged observations of explanatory and macroeconomic variables with a two-quarter lag (6 months) on the data in this study's database ($MARKETRET_{i,t-6}$; $INTRATE_{i,t-6}$). Loan volumes are lagged by one year (12 months, $LOANVOL_{i,t-12}$) due to the loan assessment cycle, a process used by banks and P2P lending platforms. During this cycle, a financial institution assesses, issues and monitors loans, deciding whether to mitigate credit risk in the event of a likely default. The typical duration of this cycle for consumer loans is about one year (Jokivuolle & Virén, 2013; Khieu et al., 2012). The current study does not use lagged observations for other control variables which generally consist of time-invariant or yearly observations. Table 1 describes all variables.

In presenting the probability of default, the specifics of P2P lending need to be considered. Due to weak regulation and insurance mechanisms, investors are faced with the high risk of not recovering their funds in the case of default (Roper, 2020). Moreover, in countries outside the US and Europe, alternative investments have a higher level of risk given that the high-risk characteristics of innovative start-ups often accompany uncertain political environments (Johan & Zhang, 2016). Therefore, some P2P lending platforms (including Mintos) offer investors a 'buyback guarantee' credit enhancement. If the loan repayment is more than 60 days late, the P2P lending platform is obligated to buy back the investment at nominal value plus accrued interest. However, the reserve funds for the 'buyback guarantee' are financed by P2P lending platforms from the same funds invested in loans. During market turmoil, lending platforms might struggle to pay their buyback guarantees, with investors sustaining substantial losses.

Thus, in this study, we use a broader definition of default probability based on loan delinquencies or 'bad loans', in line with the concepts in the existing literature for defining delinquent loans. Following Kim et al. (2018) and Wadud et al. (2020), we define delinquent loans as loans with payments 30+ days overdue and still incurring interest. Default loans are the combination of loans with default status and all charged-off loans. When combined with delinquent loans, default loans provide a broader definition of bad loans that better characterises financial distress than default loans on their own. This approach is in line with the treatment of borrowers' financial distress and insolvency by traditional financial institutions via non-performing loans (Ghosh, 2015; Louzis et al., 2012). Accordingly, this study calculates bad loans as per Eq. (2):

$$Bad\ loans_{i,t} = Default_{i,t} + Charged\ off_{i,t} + Overdue_{i,t}^{31-120\ days} \tag{2}$$

The control variables represent borrower-specific characteristics, macroeconomic conditions, demographics and country characteristics of country *i* (i.e., the country under consideration) at time *t*.[3]

We base the choice of control variables on existing traditional finance and FinTech lending studies. Prior studies on P2P lending have extensively used borrower and loan characteristics in estimating loan funding success and default (Cai et al., 2016; Galema, 2020; Serrano-Cinca et al., 2015; Wei & Lin, 2016). Emekter et al. (2015) and Serrano-Cinca et al. (2015) highlighted the importance of loan ratings (*AVERATING*) in determining loan defaults. The loan duration (*AVEDUR*) is expressed in months and provides lenders with additional information about the default risk of the loan. Loans with a shorter duration tend to signal quality by reducing asymmetric information problems and increasing the probability of selection (Menkhoff et al., 2012; Steijvers & Voordeckers, 2009). The loan amount (*AVEAMNT*) is another important indicator of solvency risk, with larger loan amounts being at greater risk to the extent that they increase default incentives.

Existing literature provides convincing evidence that general economic development significantly affects the credit risk of financial institutions (Bu et al., 2023; Dinger, 2009; Valla et al., 2006). Evidence of the positive impact of economic development on alternative financial markets is also present in studies by Khrawish et al. (2010) and Mollick (2014). Economic development and well-being are associated with GDP in empirical studies. Existing studies on crowdfunding platforms highlighted GDP growth as an important factor in the development of industry (Dushnitsky et al., 2016; Mollick, 2014). Thus, the current study includes *GDP*, the economic sentiment indicator (*ESI*), household income (*EARNINGS*) and inflation as proxy variables for economic development. Jagtiani and Lemieux

[2] Refer to Appendix A, Tables A3 and A4 for the description of loan status and loan classification.

[3] Variables are described in Table 1.

Table 1
Description of variables used in this study's analysis.

Variable	Description of variable	Source
Dependent variables		
BADLOAN	Current status of individual loan. Dummy variable equal to 1 if the loan is overdue, defaulted or buyback and 0 otherwise (current or repaid).	Mintos
SURVIVAL_TIME	Difference between the start and closing date of the loan if the loan is classified as overdue, defaulted or buyback.	Mintos
Explanatory variables		
MARKETRET	Change in daily stock market index values of country *i* at time *t* (monthly average values, decimal points).	Yahoo! Finance https://finance.yahoo.com/world-indices/
INTRATE	Interest rate on loans issued by Mintos for borrower *i* at time *t* (percentage points).	Mintos
Control variables (borrower and loan specific)		
LOANVOL	Total volume of outstanding loans issued in country *i* at time *t* (monthly, in euros, log differenced values).	Mintos
AVEAMNT	Value of individual loan (log values).	Mintos
AVEDUR	Duration of loan (in months, log values).	Mintos
AVERATING	'Mintos Rating' issued by the rating model ranging between A (1) and D (9).	Mintos
EARNINGS	Median equalised household income in country *i* at time *t* (yearly, in euros, log values).	ECB Statistical Data Warehouse http://sdw.ecb.europa.eu/
Control variables (demographic, technological, macroeconomic and financial)		
POPESTIMATE	Population of country *i* in year 2018 (log values).	OECD (2024), Population (indicator). doi: https://doi.org/10.1787/d434f82b-en
GDP	Real GDP growth (quarterly, percentage points).	OECD (2020), Quarterly GDP (indicator). doi: https://doi.org/10.1787/b86d1fc8-en
ESI	The EU economic sentiment indicator (composite measure, average = 100, log values).	Full business and consumer survey results, European Commission https://ec.europa.eu/
INFLATION	Monthly change in seasonally adjusted Consumer Price Index for all goods by country (percentage points).	OECD (2020), Inflation (CPI) (indicator). doi: https://doi.org/10.1787/eee82e6e-en
NPL	Gross non-performing loans in country *i* at time *t* (percentage of gross loans, log values).	ECB Statistical Data Warehouse http://sdw.ecb.europa.eu/
INTUSE	Percentage of individuals who have ever used the internet in country *i* at time *t* (for each year from 2015 to 2019) (yearly).	Digital economy and society, Eurostat https://ec.europa.eu/eurostat/web/digital-economy-and-society/overview
Instrumental variables		
AAR	Annualised agreed rate by credit and other institutions in country *i* at time *t* (monthly, percentage points).	ECB Statistical Data Warehouse http://sdw.ecb.europa.eu/
UNEM_RATE	Unemployment rate for each country (monthly, seasonally adjusted, percentage points).	OECD (2020), Unemployment rate (indicator). doi: https://doi.org/10.1787/b86d1fc8-en
ECBRATE	ECB deposit facility (monthly, seasonally adjusted, percentage points).	ECB Statistical Data Warehouse http://sdw.ecb.europa.eu/
STOCKIND	Average monthly stock market index of values (STOXX Europe 600) at time *t* (log values).	Yahoo! Finance https://finance.yahoo.com/world-indices/

(2018), Tanda et al. (2019) and Thakor (2020) indicated a direct relationship between traditional financial markets and alternative lending markets. Accordingly, banking sector non-performing loans (NPLs) (*NPL*) are another economy-specific variable used in this study to represent traditional financial markets.

4. Results and discussion

4.1. Description of data

This study uses the loan-book database of the Mintos platform. The study's scope covers loans issued across the multiple countries of the European Union (EU). Mintos falls into the 'P2P marketplace' category that does not create its own P2P lending platform. The Mintos marketplace platform simultaneously lists loans from multiple lending companies, known as 'loan originators', with these based in 30 countries, including 10 EU countries. Investors from 66 countries are on the marketplace platform, although Mintos does not disclose information about investor categories or origination.

The combined database comprises observations for each loan across the January 2014–December 2020 time span. The total number of unique loans included in this study is 4,485,532. The 10 EU countries included in the database are the Czech Republic, Denmark, Estonia, Finland, Latvia, Lithuania, Poland, Spain, Sweden and the United Kingdom (UK).

These countries are similar in terms of their regulatory framework and business environment. The inclusion of countries outside the EU would distort our analysis by complicating the comparisons. We include random effects identifiers in our regression models to reflect heterogeneity between the countries and the different Mintos platforms under consideration. We also run our model through bootstrap samples based on the country of origination to adjust for a possible sampling bias. Table 1 lists and describes variables used in this study.

4.2. Descriptive statistics

Table 2 reports the descriptive statistics for the variables used in the study's analyses. The combined mean value of issued monthly loans by Mintos in each country was €4.86 million, with an average of €510,000 of monthly issued loans classified as bad loans, bringing the average probability of default to 11.17 %.[4]

The mean value for the interest rate for loans in the sample is 13.92 %, ranging between the lower quartile of 6.09 % and the upper quartile of 14.45 %. The average interest rate and net return to investors are considered substantially different in the P2P lending market. The platforms, with their diversification and assessment of loans, try to offset high loan defaults with a higher interest rate spread. In this regard, the proper estimation of the probability of default is vital for P2P lending platforms to meet the returns required by investors.

The average amount of loans issued is €2291, with an average duration of 38.72 months. The average borrower rating of loans[5] is 5.58, which falls between 'B-' and 'C+'. The mean average income of borrowers is €15,025, with a standard deviation of €5207, while the median income is lower at €11,540.

This study presents a correlation matrix for dependent, explanatory and control variables, in Table 3, which shows low values for Pearson's correlation coefficients for most variable pairs with few exceptions. Most importantly, we do not observe any high correlation coefficients for two variables of interest: stock market returns (*MARKETRET*) and interest rate (*INTRATE*).

High correlation coefficients are observed between variables not used in the same model. For example, the correlation coefficient between *NPL* and *AVERATING* is 0.6196, indicating a strong positive correlation. However, *AVERATING* is not used as an independent variable in our regression models. Instead, we conduct additional analysis by creating subsamples of loans based on their credit ratings.

Another significantly high correlation coefficient is observed between *ESI* and *ECBRATE*. The high correlation level between these variables is predictable as *ECBRATE* closely follows economic sentiment. However, we only use *ECBRATE* as an instrumental variable for the test of endogeneity.

We also carried out the test for potential multicollinearity using variance inflation factor (VIF) values. The VIF values for the variables used in our study's models are reported in Table 4. We observe that the reported values are well below 10 (ranging from 1.03 to 4.35), with all tolerance statistics ranging above 0.2. This confirms the general absence of multicollinearity in the models.

4.3. Results of baseline regression analysis

This study explores the impact of borrower- and country-specific factors on the probability of loan default. Stock market return and interest rate are emphasised as the main explanatory variables. Table 5 shows the estimation results for baseline regression with the probability of default as the dependent variable. The results are based on the logit estimation method with the constant, as well as year and individual (loan originator) effects. The sample comprises valid observations ($N = 4,485,532$) across all models.

The fit of the models shown in Table 3, based on overall *R*-squared (R^2) values, varies between 0.0547 and 0.1188. These values indicate that independent variables collectively explain around 5.47–11.88 % of changes in the dependent variable. The low pseudo-*R*-squared values observed in Table 5 could be attributed to several factors, such as model simplicity and non-linear relationships. In our study, we attribute the low pseudo-*R*-squared values to non-linear relationships and inherent variability in the target variable as we are analysing a database of several million observations from different countries. We address issues related to non-linear relationships and inherent variability through robustness tests and in the additional analyses section (Section 4.5). However, while low pseudo-*R*-squared values may seem modest, they should be interpreted alongside other contextual information and domain knowledge to assess the overall model fit and its practical significance.

Table 5, Models (1)–(4) depict the probability of default as a function of stock market return (*MARKETRET*) and control variables. Stock market return (*MARKETRET*) is the monthly average value of daily changes in stock market indices for the respective countries in the database. Model (1) results indicate that the coefficient of *MARKETRET* is significant and positive at the 0.0977 level. The coefficients for *MARKETRET* are robust with the inclusion of loan- and country-specific control variables in Models (2)–(4). The coefficients for *MARKETRET* are significantly positive, varying from 0.0873 to 0.1690. At the same time, based on the odds ratio, a default event is more likely to occur when market returns increase (i.e., the odds ratio is higher than 1).

The findings for *MARKETRET* reveal important information on the interaction between the traditional financial market and the P2P lending market. The existing literature is mostly lacking in this regard, with only one study known to empirically examine the interaction between these markets. Jagtiani and Lemieux (2018) analysed the penetration of P2P lending loans in various geographical areas of the United States (US). They indicated that P2P consumer lending penetrated areas with relatively underserved banking markets and in which the local economy was not performing well. Moreover, the current study can refer to several studies that explored the relationship between traditional and alternative lending markets. Buchak et al. (2018) and Thakor (2020) highlighted that alternative investments prosper when conventional lending channels are not developed or traditional banks face regulatory

[4] This percentage represents the average ratio reported for each country and platform.

[5] Borrower ratings range between A (1) and D (9).

Table 2
Descriptive statistics for variables included in regression analysis.

Variables	Mean	Median	Standard deviation	Minimum	Maximum	Lower Quartile	Upper Quartile
Dependent variable							
BADLOAN (€ thousands)	510.7155	551.1167	227.0903	60.8285	620.6151	67.0805	510.7155
Explanatory variables							
MARKETRET	−0.0019	0.0000	0.0182	−13.1342	8.4432	−0.0355	0.0450
INTRATE (%)	13.9237	13.4872	4.9328	5.7047	25.1539	6.0931	14.4529
Borrower- and loan-specific controls							
LOANVOL (€ thousands)	4863.3630	4950.9000	1725.5534	23.1816	8128.2000	40.5856	6028.9000
AVEAMNT	2291.4355	2264.5570	1259.7675	110.0000	200,000.0000	123.9360	4904.1910
AVEDUR (months)	38.7239	43.2911	13.7434	1.0000	240.0000	8.0674	56.1027
AVERATING	5.5838	4.0000	1.2704	1.0000	9.0000	3.5343	7.9241
EARNINGS	15,025.5100	11,540.3500	5207.7520	4085.5100	37,004.4100	11,540.3500	16,413.3000
Demographic, technological, macroeconomic and financial controls							
GDP (%)	0.6113	0.6777	0.8366	−3.6483	2.2806	−3.3093	1.9374
ESI	99.0694	99.7000	8.6299	46.9000	115.6000	63.9000	111.3000
INFLATION (%)	0.7471	0.2600	4.6397	−0.1100	11.7600	0.2200	8.3800
POPESTIMATE (millions)	2.9403	1.3209	4.7396	1.3209	5.5181	1.3209	5.5181
NPL	3.0482	3.1000	1.7095	1.0000	20.7000	1.2000	8.1000
INTUSE (% of population)	88.1075	90.0000	6.2216	46.0000	99.0000	76.0000	97.0000
Instrumental variables							
AAR (%)	11.7562	9.6000	6.6322	2.8600	33.2400	2.9400	23.0600
ECBRATE (%)	−0.4030	−0.4000	0.0213	−0.5000	3.2500	−0.5000	3.0400
STOCKIND	388.2007	389.5700	94.2875	329.7300	424.3600	357.7600	400.1000
UNEM_RATE (%)	7.1314	6.4000	3.8373	2.9000	20.5000	3.9000	8.1000

Note: Table 2 reports the descriptive statistics for variables included in the regression analysis. The variables include platform-specific, economic, demographic, technological and political characteristics. Descriptions of variables are provided in Table 1.

constraints.

In this study, we observe that, with an increase in stock market returns, delinquencies in the P2P lending market also increase. As highlighted in Section 2, stock market conditions could bilaterally influence the default risk in bank credit markets (Ericsson et al., 2009; Norden & Weber, 2009; Yoon et al., 2019). The current study provides the first evidence of the same relationship in the P2P lending market.

Table 5 also reports the results of estimating the coefficients based on Eq. (1) with the interest rate (*INTRATE*) as the main explanatory variable. Table 5, Models (5)–(8) use the same set of control variables as Models (1)–(4). The interest rate is significantly positive across all four models, varying between 0.1568 and 0.4925 (i.e., the odds ratio is higher than 1), indicating that the higher the interest rate, the higher the likelihood of loan default. Delinquency rates, represented by bad loans, are found to be significantly and positively affected by interest rates.

These interest rate findings are consistent with those in earlier studies by Nigmonov et al. (2022) and Serrano-Cinca et al. (2015). However, these earlier studies investigated this issue using a single US platform (LendingClub). In our study, the heterogeneity between the panels is more pronounced in the European cross-country database. Thus, the current study provides robust results for country-level interest rates rather than for rates set for individual loans. The results are in line with the findings of similar empirical studies in traditional financial markets (Beck, 2013; Espinoza & Prasad, 2010) but contradict the findings of studies by Ghosh (2015); Goel and Hasan (2011); and Jakubík (2006). The impacts of *INTRATE* and *MARKETRET* require further robustness tests as discussed in Section 4.4.

Table 5 presents convincing results in terms of control variables. The coefficients of *LOANVOL* are consistent across Table 5, Models (1)–(8), being significantly negative and varying between −0.5047 and − 0.3940. This means that a 1 % increase in P2P loan volume decreases the probability of default in the range of 0.3940–0.5047 %. Each country's population is found to have significant positive coefficients ranging from 0.1959 to 0.5102. This finding indicates that larger countries, in terms of population, have higher default rates among P2P loan borrowers. The individual loan amounts (*AVEAMNT*) are found to be positively associated with the probability of default. On the other hand, the coefficients of average loan duration (*AVEDUR*) yield significant negative coefficients, indicating that loans with a longer duration have a lower probability of default.

The baseline regression analysis indicates important findings in terms of the impact of macroeconomic variables on the probability of default in the P2P lending market. The coefficients of *GDP* are insignificant across all models. This study extends the economic development indicators by using the economic sentiment indicator (*ESI*) as a proxy, although the coefficients of *ESI* are inconsistent and insignificant across all models in Table 5. Our study includes an additional indicator of financial market performance, namely, banking sector non-performing loans (*NPL*) for each country, with *NPL* having insignificant coefficients in Table 5, Models (4)–(8).

Table 3
Correlation matrix.

	BADLOAN	*LOANVOL*	*INTRATE*	*MARKETRET*	*INCOME*	*DTI*	*AVEAMNT*	*AVEDUR*	*AVERATING*
BADLOAN	1.0000								
LOANVOL	0.2961***	1.0000							
INTRATE	0.0552***	−0.0923***	1.0000						
MARKETRET	0.0167***	−0.0346***	−0.0110***	1.0000					
AVEINCOME	−0.1319***	−0.1579***	−0.0421***	0.0159***	1.0000				
DTI	−0.3105***	−0.3550***	0.0613***	−0.0319***	−0.3047***	1.0000			
AVEAMNT	0.3549***	0.2751***	−0.1346***	−0.0355***	−0.3385***	0.3251***	1.0000		
AVEDUR	0.4959***	0.4211***	0.1712***	−0.0796***	−0.4591***	0.0869***	0.5262***	1.0000	
AVERATING	−0.2218***	−0.4130***	0.5791***	−0.0064***	0.1851***	0.0344***	−0.1573***	−0.2374***	1.0000
INFLATION	0.0662***	0.0805*	−0.0106***	−0.0165***	0.0232***	0.0052***	0.0425***	0.0405***	0.0256***
GDP	0.0289***	0.0588***	−0.0116***	0.0060***	−0.2062***	0.0606***	−0.1019***	0.0098***	−0.1066***
AAR	−0.2392***	−0.0562***	−0.3757***	0.0020***	0.0972***	−0.0619***	−0.4258***	−0.4567***	−0.3494***
ESI	0.1454***	0.0041***	0.4040***	0.0019***	−0.3502***	−0.0016***	0.0026***	0.2798***	0.1471***
ECBRATE	0.3005***	0.1645***	0.3821***	−0.0167***	−0.3918***	−0.0429***	0.1120***	0.3762***	0.0692***
INTUSE	0.5501***	0.5149	−0.0114***	0.0189***	−0.3324***	−0.0188***	0.5975	0.6670***	−0.4251***
INDEX	0.2092***	0.4843***	0.4678***	0.0407***	−0.2593***	−0.2016***	−0.0016***	0.4308***	0.2678***
NPL	−0.3892**	−0.3772***	0.2756***	0.011***	−0.0458***	0.1583***	−0.3318***	−0.2705***	**0.6196*****
POPESTIMATE	−0.0355***	−0.0486***	0.3095***	0.0258***	−0.1022***	−0.0536***	−0.1216**	0.1017***	0.3149***
EARNINGS	0.5430***	0.5445***	0.0823***	0.2800*	−0.3366***	−0.0259***	0.5986***	0.5033***	−0.3341***
UNEM_RATE	0.3000***	0.4843***	−0.2373***	0.0298***	−0.1430***	−0.3524***	−0.0709***	0.1427***	−0.5213***

	INFLATION	*GDP*	*AAR*	*ESI*	*ECBRATE*	*INTUSE*	*INDEX*	*NPL*	*POPESTIMATE*	*EARNINGS*
INFLATION	1.0000									
GDP	0.0120***	1.0000								
AAR	−0.0231***	0.3533***	1.0000							
ESI	−0.0165***	0.5475***	−0.0122	1.0000						
ECBRATE	−0.0287***	0.5051***	−0.0096***	**0.9234*****	1.0000					
INTUSE	0.0439***	−0.0349***	−0.4999***	0.0413	0.1753***	1.0000				
INDEX	0.0451***	−0.0370***	−0.5625***	0.1023	0.2307***	0.4887***	1.0000			
NPL	0.1017***	0.0921***	−0.2913***	0.3862	0.2969***	0.0423***	0.1164***	1.0000		
POPESTIMATE	0.1171***	0.1854***	0.2406***	−0.0421	−0.0663***	0.2433***	0.1979***	0.4314***	1.0000	
EARNINGS	0.0019***	0.0457***	−0.3339***	0.2473	0.2489***	0.4298***	0.4513***	0.4352***	0.2189***	1.0000
UNEM_RATE	0.0020***	0.0457***	0.0705**	0.1492	0.0085***	−0.0790***	−0.0693***	0.3700***	−0.2641***	−0.1452***

Note: Table 3 reports Pearson's correlation coefficients between the variables employed in regression analyses of this study (*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$). Descriptions of variables are provided in Table 1.

Table 4
Variance inflation factor (VIF) results from regression model.

Variable	VIF	1/VIF
ESI	4.35	0.2299
NPL	3.82	0.2618
INTRATE	3.01	0.3322
MARKETRET	2.49	0.4016
LOANVOL	1.87	0.5348
AVERATING	1.71	0.5848
AVEDUR	1.63	0.6135
EARNINGS	1.51	0.6623
GDP	1.38	0.7246
POPESTIMATE	1.22	0.8197
INTUSE	1.15	0.8696
AVEAMNT	1.06	0.9434
INFLATION	1.03	0.9709
Mean VIF	**2.02**	

Note: Table 4 Variance inflation factor (VIF) for the variables used in the regression models.

4.4. Results for instrumental variable estimation

Our baseline regression results may be subject to the problem of endogeneity. Reverse causality between the two main explanatory variables (*INTRATE* and *MARKETRET*) and other variables in the model cannot be ruled out. In fact, growth in FinTech tends to reduce the capital available to more efficient firms through FinTech-induced competition in the lending market (Xie & Zhu, 2022). Moreover, competition from P2P lending may force traditional banks to open their own platforms or to take over existing ones to prevent significant erosion of lending volume (Thakor, 2020). Thus, reverse causality and/or simultaneity, as well as other sources of endogeneity bias, could inflate the estimated parameters, making them the upper bound of the true relationship.

The regression models in the current study account for year and individual effects, in line with a broad range of country-specific indicators. Nevertheless, the baseline regression model may have time-varying omitted variables correlated with the explanatory variables. Due to several macroeconomic and loan-specific indicators, explanatory variables *INTRATE* and *MARKETRET* might be driven by control variables, such as *LOANVOL* and *GDP*, respectively.

The instrumental variables (IVs) approach is used to solve the inherent endogeneity problem in the current study's baseline regression models. In line with Hakura and Cosimano (2011), in the first-stage regression, unemployment rate (*UNEM_RATE*) and European Central Bank (ECB) deposit facility rate (*ECBRATE*) are used as instrumental variables for stock market return. In a separate model, the stock market index (*STOCKIND*) and annualised agreed rate (*AAR*) are used as instrumental variables for the interest rate. The estimates consequently define the explanatory variable, with the second-stage regression then employing *BADLOAN* as the dependent variable, with estimated stock market return and interest rate as explanatory variables. In general terms, this model allows estimation of the effect of determinants on *BADLOAN* in a context which could not be accomplished by our baseline regression analysis.

In this case, our study's model may be exposed to outliers in the data as loan originators are from certain specific countries (as reported in Table A2). For instrumental variables, the classical parametric estimators can become arbitrarily biased (Freue et al., 2013; Ronchetti & Trojani, 2001). If a small model misspecification is present, results may be biased by a cluster of atypical observations (Andrews et al., 2019; Young, 2022).

In addition, with binary dependent variables (taking values 0 or 1), the linearity and distributional assumptions do not hold. In such cases, using a linear model to estimate the relationship between the dependent and independent variables might not be appropriate as predicted values could fall outside the values of 0 or 1. To address this concern about the stability of estimators and tests, we estimate two-stage probit regression with instrumental variables using maximum likelihood estimation (MLE). Following the propositions of Amemiya (1978) and Rivers and Vuong (1988), this method of estimation resolves problems related to distributional assumptions and the predicted variables falling outside the values of 0 and 1.

Table 6 presents the two-stage probit regression results with instrumental variables for the interest rate and stock market return. As shown in Table 6, all instrumental variables yield significant coefficients, indicating that these variables are statistically significant predictors of interest rate and stock market return. In the second stage, Models (1) and (2) show that both interest rate and stock market return retain their significantly positive coefficients. Thus, the interest rate and stock market return have significant positive coefficients in the baseline models and are statistically significant in the models with instrumental variables. Instrument diagnostic tests also favour the use of two models with their respective instrumental variables as shown in Table 6.

This study, due to its high dimensionality, recognises some limitations when using the baseline regression approach. The various explanatory and control variables employed by the study have been extensively used in the traditional financial literature. However, few prior studies on P2P lending have employed this wide combination of variables. Accordingly, our study may be prone to the multicollinearity problem which could potentially blur the results. To further strengthen the regression model, the study implements a type of machine learning process, known as the least absolute shrinkage and selection operator (LASSO). This method provides a more robust analysis that allows the important variables to be found in a large set of potential determinants (Belloni et al., 2016; Tibshirani, 1996). The method shrinks regression coefficients by penalising their magnitude and provides a narrow set of important variables,

Table 5
Impact of stock market returns and interest rates on probability of default.

Panel A: Impact of stock market returns on probability of default

	Model 1		Model 2		Model 3		Model 4	
Variables	Dependent variable (DV) = *BADLOAN*							
	Coefficient	Odds ratio	Coefficient	Odds ratio	Coefficient	Odds ratio	Coefficient	Odds ratio
MARKETRET	0.0977***	1.1540	0.1226***	1.1939	0.1690***	1.2140	0.0873***	1.1239
	(0.0011)		(0.0023)		(0.0025)		(0.0011)	
POPESTIMATE	0.1689***	1.2159	0.2146***	1.2226	0.2102***	1.1761	0.1872***	1.0948
	(0.0269)		(0.0135)		(0.0113)		(0.0117)	
LOANVOL	−0.5036***	0.7667	−0.4251***	0.7938	−0.3885***	0.7117	−0.3874***	0.7137
	(0.0480)		(0.0465)		(0.0492)		(0.0492)	
AVEAMNT	0.5888***	1.4741	0.5692***	1.3040	0.5934***	1.4059	0.5835***	1.4109
	(0.0707)		(0.0612)		(0.0646)		(0.0703)	
AVEDUR	−0.0012***	0.9175	−0.0012***	0.9131	−0.0011***	0.9667	−0.0010***	0.9219
	(0.0001)		(0.0001)		(0.0002)		(0.0002)	
AVERATING			0.0697*	1.0567	0.0492	1.0741	0.0429	1.0738
			(0.0410)		(0.0422)		(0.0459)	
EARNINGS			0.0594	1.0249	0.1487	1.1756	0.1592	1.2669
			(0.0894)		(0.1052)		(0.1256)	
GDP					0.0312	1.1669	0.0074	1.0756
					(0.0962)		(0.0967)	
ESI					−0.0029	0.8883	0.0003	1.0062
					(0.0107)		(0.0110)	
INFLATION					−0.0040	0.9174	−0.0021	0.9883
					(0.0249)		(0.0249)	
NPL							0.0441	1.0747
							(0.0367)	
INTUSE							0.0246**	1.1740
							(0.0112)	
Constant, Yr. & Ind. Effects	Yes		Yes		Yes		Yes	
LR chi^2	200,939.7224		122,487.2334		154,300.0023		297,073.1992	
Prob>chi^2	0.0000		0.0000		0.0000		0.0000	
Pseudo-*R*-squared	0.0799		0.0918		0.1006		0.1075	
N	4,485,532		4,485,532		4,485,532		4,485,532	

Panel B: Impact of interest rates on probability of default

	Model 5		Model 6		Model 7		Model 8	
Variables	DV=*BADLOAN*							
	Coefficient	Odds ratio	Coefficient	Odds ratio	Coefficient	Odds ratio	Coefficient	Odds ratio
INTRATE	0.1568***	1.3295	0.3485***	1.2705	0.4899***	1.2258	0.4925***	1.3340
	(0.0010)		(0.0029)		(0.0033)		(0.0013)	
POPESTIMATE	0.2701***	1.0342	0.1959***	1.0335	0.2625***	1.0304	0.2574***	1.0381
	(0.0959)		(0.0758)		(0.0840)		(0.0856)	
LOANVOL	−0.5047***	0.8850	−0.4444***	0.8643	−0.3940***	0.8767	−0.3960***	0.8790
	(0.0508)		(0.0478)		(0.0509)		(0.0512)	
AVEAMNT	0.5516***	1.2382	0.5146***	1.2569	0.5049***	1.2289	0.5170***	1.2318
	(0.0742)		(0.0646)		(0.0689)		(0.0740)	
AVEDUR	−0.0009***	0.9768	−0.0012***	0.9393	−0.0009***	0.9463	−0.0009***	0.9463
	(0.0002)		(0.0002)		(0.0002)		(0.0002)	
AVERATING			0.0382	1.0203	−0.0068	0.9392	−0.0144	0.9591
			(0.0464)		(0.0486)		(0.0533)	
EARNINGS			0.1055	1.2754	0.2322**	1.3302	0.2653**	1.3451
			(0.0910)		(0.1058)		(0.1270)	
GDP					0.0902	1.1673	0.0907	1.1678
					(0.0973)		(0.0980)	
ESI					−0.0100	0.9103	−0.0108	0.9108
					(0.0108)		(0.0112)	
INFLATION					0.0170	1.0015	0.0164	1.0011
					(0.0255)		(0.0256)	
NPL							0.0190	1.0138
							(0.0377)	
INTUSE							0.0017	1.0241
							(0.0120)	
Constant, Yr. & Ind. Effects	Yes		Yes		Yes		Yes	
LR chi^2	150,972.3953		327,478.2578		140,410.3251		127,671.9178	

(*continued on next page*)

Table 5 (*continued*)

Panel B: Impact of interest rates on probability of default								
	Model 5		Model 6		Model 7		Model 8	
Variables	DV=*BADLOAN*							
	Coefficient	Odds ratio	Coefficient	Odds ratio	Coefficient	Odds ratio	Coefficient	Odds ratio
Prob>chi^2	0.0000		0.0000		0.0000		0.0000	
Pseudo-*R*-squared	0.0547		0.1188		0.0818		0.1054	
N	4,485,532		4,485,532		4,485,532		4,485,532	

Note: Table 5 reports the results for logit estimation presenting the effect of stock market returns and interest rates on the probability of default with control variables. Estimations are based on Eq. (1). The estimation model employs *BADLOAN* as the dependent variable. All model specifications employ robust standard errors of coefficients in parentheses; $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6
Probit model with endogenous regressors: Impact of stock market returns and interest rates on probability of default.

	Model 1: Stock market returns and probability of default		Model 2: Interest rate and probability of default	
	1st Stage	2nd Stage	1st Stage	2nd Stage
	DV = *MARKETRET*	DV=*BADLOAN*	DV = *AVEINTRATE*	DV=*BADLOAN*
MARKETRET/ INTRATE		0.4738***		3.3582***
		(0.1052)		(0.7955)
UNEM_RATE	1.3878***			
	(0.4351)			
ECBRATE	−4.9302***			
	(1.0906)			
STOCKIND			0.0837***	
			(0.0135)	
AAR			0.0331	
			(0.0321)	
POPESTIMATE	−3.0043***	1.4613***	−0.0166	0.2910***
	(0.2500)	(0.2964)	(0.0211)	(0.1028)
LOANVOL	−0.0529	−0.3539***	−0.0953***	0.0373
	(0.1222)	(0.0508)	(0.0090)	(0.0945)
AVEAMNT	−0.4916***	0.8061***	0.1087***	0.2575**
	(0.1781)	(0.1008)	(0.0164)	(0.1209)
AVEDUR	−0.0019***	0.0001	−0.0003***	0.0006*
	(0.0004)	(0.0003)	(0.0000)	(0.0003)
AVERATING	−0.0573	0.0722	0.1589***	−0.4715***
	(0.1251)	(0.0663)	(0.0115)	(0.1362)
EARNINGS	1.3394***	−0.4613*	−0.0236	0.3180**
	(0.3481)	(0.2763)	(0.0378)	(0.1515)
GDP	0.9898***	−0.4580***	−0.0113	0.0212
	(0.2404)	(0.1699)	(0.0225)	(0.1088)
ESI	−0.0806***	0.0448***	0.0074***	−0.0099
	(0.0237)	(0.0140)	(0.0022)	(0.0123)
INFLATION	0.1341*	−0.0357	−0.0170***	0.0923***
	(0.0694)	(0.0406)	(0.0064)	(0.0335)
NPL	−0.2139**	0.1287**	0.0292***	−0.0435
	(0.0924)	(0.0500)	(0.0085)	(0.0491)
INTUSE	−0.1481***	0.0907***	0.0263***	−0.0688**
	(0.0303)	(0.0199)	(0.0028)	(0.0281)
Constant, Yr. & Ind. effects	Yes	Yes	Yes	Yes
N		4,485,532		4,485,532
Instrument diagnostics tests:				
Test of endogeneity: Wald test of endogeneity				
		5.1091***		6.0114***

Note: Table 6 presents probit model results with endogenous regressors of interest rate and stock market returns using maximum likelihood estimation (MLE). All model specifications employ robust standard errors in parentheses; $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

making the results easier to interpret and resolving the problem of multicollinearity (Meinshausen & Yu, 2009).

Using the LASSO method, our study could select a set of variables that may be more important in determining the default risk among Mintos P2P loans. Table 7, Panel A presents the results of the LASSO model for the impact of stock market returns and interest rates on the probability of default (with control variables). Three types of the LASSO method (cross-validation, plug-in and adaptive) are used to estimate effects for the potential independent and control variables to be included in the model. We find that the coefficients of the independent and control variables are both similar in the baseline regression. Most explanatory variables also hold their respective coefficient signs in the LASSO model.

Table 7
Impact of stock market returns and interest rates on probability of default: LASSO selection model.

Panel A: LASSO selection models

Variables	Model 1		Model 2		Model 3	
	DV=*BADLOAN*					
MARKETRET	0.1926		0.1028		0.1584	
INTRATE		0.5669		0.1225		0.5327
POPESTIMATE	−0.0221	−0.0260	−0.0013	−0.0019	−0.0230	−0.0255
LOANVOL	0.1608	0.1366	X	X	0.1594	0.1084
AVEAMNT	0.7239	0.7854	X	X	0.7475	0.7548
AVEDUR	−0.2515	0.0889	−0.0467	X	−0.2148	X
AVERATING	0.0833	0.0813	X	0.0029	0.0839	0.0720
EARNINGS	0.1247	0.1168	0.0294	0.0289	0.1244	0.1122
GDP	0.0313	0.0332	X	X	0.0319	0.0300
ESI	−0.2488	−1.4348	0.5134	0.3994	X	−0.9603
INFLATION	−0.1873	−0.2253	−0.0270	−0.0321	−0.1907	−0.2188
NPL	2.1033	2.4509	0.3107	0.3494	2.1560	2.3734
INTUSE	X	0.1258	−0.1698	−0.1346	X	X
Constant, Yr. & Individual Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4,485,532	4,485,532	4,485,532	4,485,532	4,485,532	4,485,532

Panel B: LASSO inference models

Variables	Model 1		Model 2		Model 3	
	DV=*BADLOAN*					
MARKETRET	0.4669***		0.4682***		0.5845*	
	(0.0315)		(0.0314)		(0.3136)	
INTRATE		0.5628***		0.5607***		0.5628***
		(0.0668)		(0.0670)		(0.0668)
Constant, Yr. & Individual Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4,485,532	4,485,532	4,485,532	4,485,532	4,485,532	4,485,532

Note: Table 7 presents the LASSO method results for the impact of stock market returns and interest rates on the probability of default with control variables. The LASSO method is used to estimate effects for potential independent and control variables included in the model. Model (1) uses the adaptive LASSO model; Model (2) uses the LASSO model with the cross-validation (CV) method; and Model (3) employs the LASSO plug-in method. Omitted variables from selection by the LASSO method are denoted as (X). Panel B presents LASSO inference models based on double-selection LASSO logistic regression.

It should also be noted that the LASSO models are inherently selection models. This group of models selects covariates and estimates coefficients without providing standard errors (Bühlmann & van de Geer, 2011). However, several modified versions of the LASSO model allow the derivation of standard errors of estimates. Specifically, the double-selection LASSO method uses selected control variables in the inference model to estimate effects for variables of interest (Belloni et al., 2014a). We present the results of the double-selection LASSO method in Table 7, Panel B, in which the selected control variables for the variables of interest are included in the model. It should be noted that the double-selection LASSO method does not provide estimates of the coefficients of the control variables or their standard errors (Belloni et al., 2014b). However, the estimation results for the two variables of interest, that is, stock market return and interest rate, are similar to the baseline regression findings.

It may also be argued that utilisation of selected instrumental variables, such as the unemployment rate, are not well justified. As previously highlighted, the main cause for endogeneity is reverse causality. This argument may be the less likely one as loan delinquency is very unlikely to change economy-wide factors, such as market returns. In fact, some instrumental variables, such as the unemployment rate, are direct determinants of loan delinquency, leading to a violation of the exclusion restrictions. We address this major challenge by employing another recent extension of the LASSO method that allows for instrumental variable regression. This method identifies valid instrumental variables and deals with endogenous covariates more efficiently (Chernozhukov et al., 2015).

We apply partialling-out LASSO instrumental variable regression to the estimate coefficients and standard errors for the variables of interest, both exogenous and endogenous. Two LASSO models are used to select from potential control variables and instruments: the first selects variables from all control variables, while the second chooses from country-specific variables only. In doing so, we obtain additional evidence that the instrumental variable regressions reported in Table 6 do not suffer from the model selection problem.

The selected variables and coefficients for interest rate and stock market volatility are reported in Table 8. The LASSO method selects 8–9 control variables used in the original baseline regression (reported in Table 5) and 1–2 instruments from the four potential variables used in the endogeneity correction (reported in Table 6). Both the explanatory variables, namely, *MARKETRET* and *INTRATE*, have coefficient signs and significance levels similar to those reported in the baseline regression. Therefore, we can conclude that the selection of both control and instrumental variables is well justified and does not significantly affect the impact of explanatory variables on default risk.

The next potential shortcoming of our analysis is that the sample comprises heterogeneously distributed observations across loan originators. For instance, listings are overrepresented by loan originators based in Latvia, Poland and Spain. This feature of the database potentially creates complications related to sample selection bias. Another potential problem with the sample database is the

Table 8
Impact of stock market returns and interest rates on probability of default: LASSO models with instrumental variables.

	Model 1	Model 2
Panel A: LASSO selection model with all control variables		
	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET	0.1330***	
	(0.0219)	
INTRATE		2.4910***
		(0.6759)
Constant, Yr. & Ind. effects	Yes	Yes
Number of selected control variables (out of 11)	8	8
Selected instruments	*STOCKIND, AAR*	*STOCKIND, UNEM_RATE*
N	4,485,532	4,485,532
Panel B: LASSO selection model with only country-specific variables from which to choose		
	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET	0.1639***	
	(0.0254)	
INTRATE		2.6518***
		(0.5752)
Constant, Yr. & Ind. effects	Yes	Yes
Number of selected control variables (out of 11)	8	9
Selected instruments	*STOCKIND*	*ECBRATE, UNEM_RATE*
N	4,485,532	4,485,532

Note: Table 8 presents the results of the LASSO instrumental variables linear regression model for the probability of default (PD). The cross-fit partialling-out method is used to estimate the effects for potential control variables and instruments to be included in the model. Panel A uses all control variables for the LASSO model. Panel B includes loan- and borrower-specific variables and uses country-specific variables (economic, demographic, etc.) for the LASSO model. Both Panels A and B use all four instrumental variables for selection by the LASSO method. Refer to Table 1 for descriptions of variables. All model specifications employ robust standard errors in parentheses; $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

heterogeneous distribution of loans with a clear ending resolution (i.e., resolved loans) which account for around 92 % of the database. Unresolved loans are mainly concentrated around the later dates of the database. This concentration of unresolved loans potentially creates a misrepresentation of the sample selection as loans included in the earlier periods might be more likely to be defaulted or buyback loans.

To address these issues, we employ three different procedures. Firstly, we use a random bootstrap sampling[6] technique to obtain robust estimates of the relevant coefficients. This method reduces the sampling bias and warrants that our estimates are not affected by the under-weighting or over-weighting of a certain group of observations (Chernick & LaBudde, 2014; Tibshirani & Efron, 1993). Secondly, we apply Heckman's (1979) selection model for sample selection to the binary regression model. A binary dependent variable equal to 1 is used if the loan is 'resolved' (i.e., has a clear outcome), and 0 otherwise. The sample selection is instrumentalised with loan-specific variables, including the loan rating. Thirdly, we run an ordered probit model in which the dependent variable is the status of the unresolved loan ('current', 'in grace period' or 'a late loan[7]'). Accordingly, the dependent variable *(LOANSTATUS)* takes one of the six values, with the regression sample consisting of only unresolved loans.

Table 9 reports the results of the probit regression after controlling for the above techniques to address selection bias. We observe that the results are identical to the baseline regression reported in Table 5 in terms of coefficient signs and significance. The results are generally robust to three specifications and remain similar to the ones reported in the baseline regression. We conclude that the detected impacts of *MARKETRET* and *INTRATE* are almost not affected by the selection mechanisms used to construct the sample.

4.5. Additional analyses

In recognising that interest rates are calculated for aggregated individual loans, we acknowledge that individual personality characteristics may affect loan defaults. Specifically, if the ratio of loans to borrowers in a country is low, the average-level data may not be useful for this identification and may cause sampling bias. Therefore, we further analyse regression models within the context of loan volume by dividing the sample into two groups based on the median loan volume. The first subsample comprises observations with lower than median loan volume while the second subsample comprises observations with higher than median loan volume. In aggregating the data set by individual countries, we find that variations in terms of issued loans are very diverse.

Table 10, Panel A provides regression results for these two models with respective two-sample coefficient tests. We find

[6] Stratified bootstrap samples based on loan originators and each year under consideration. Bootstrap sampling replications are conducted 1000 times for each of the regression estimations.

[7] Late loans are further classified as: '1–15 days' late', '16–30 days' late', '31–60 days' late' and '60+ days' late'.

Table 9
Impact of stock market returns and interest rates on probability of default: Bootstrap sampling, Heckman's selection model and ordered probit regression.

	Model 1	Model 2
Panel A: Bootstrap sampling		
	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET	0.1856***	
	(0.0110)	
INTRATE		0.3537***
		(0.0163)
Controls	Yes	Yes
Constant	Yes	Yes
Yr. & Ind. effects	Yes	Yes
Pseudo-R-squared	0.0971	0.0607
N	4,485,532	4,485,532
Panel B: Heckman's selection model		
	DV=*BADLOAN*	**DV=*BADLOAN***
MARKETRET	0.1618***	
	(0.0026)	
INTRATE		0.4999***
		(0.0034)
Controls	Yes	Yes
Constant	Yes	Yes
Yr. & Ind. effects	Yes	Yes
Pseudo-R-squared	0.1013	0.1199
N	4,485,532	4,485,532
Panel C: Ordered probit regression		
	DV = *LOANSTATUS*	**DV = *LOANSTATUS***
MARKETRET	0.1931***	
	(0.0340)	
INTRATE		0.1050***
		(0.0184)
Controls	Yes	Yes
Constant	Yes	Yes
Yr. & Ind. effects	Yes	Yes
Pseudo-R-squared	0.0670	0.0802
N	4,485,532	4,485,532

Note: Table 9 presents the probit regression results for three panels with corrections for sampling bias. Panel A reports the results after the application of bootstrap sampling with stratified sampling based on country and year. Panel B reports the results after the application of the Heckman selection model. Panel C reports the results for ordered probit regression analysis for the loan status (*LOANSTATUS*) with the sample consisting of only unresolved loans. The dependent variable is an ordered dependent variable which takes one of the six values (current, in grace period, 1–15 days' late, 16–30 days' late, 31–60 days' late and 60+ days' late). Refer to Table 1 for descriptions of variables. All model specifications employ robust standard errors in parentheses; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

qualitatively similar results when using the effects of both stock market return and interest rate on the probability of default. The explanatory variables behave similarly with no significant difference between the subsamples. The two-sample coefficient t-tests indicate that the differences between the coefficients of the two groups are insignificant, with Chow (1960) test chi-squared statistic values of 0.71 and 0.68 for market return and interest rate, respectively. The impacts of both interest rate and stock market return do not significantly differ with changes in the volume of issued loans, thus eliminating sampling bias based on the effects of personality characteristics.

We also notice that our sample period roughly covers two different periods of market interest rate conditions. Following the GFC, the ECB kept baseline interest rates near zero for a prolonged period. Specifically, from 16 March 2016, the ECB kept the interest rate on the main refinancing operations at zero (European Central Bank (ECB), 2021). Owing to this policy, our sample database contains observations for periods when the euro area base interest rate was zero, during which little volatility was observed in financial markets. Therefore, we draw separate subsamples based on a proxy for the market interest rate of the annualised agreed rate (*AAR*) reported by credit organisations and other institutions. As shown in Table 10, Panel B, loan defaults are more prone to the effects of both interest rates and market returns in a higher interest rate environment when the *AAR* is used as a proxy. This result has important implications as interest rates are slated to increase due to the inflationary pressure faced by central banks (Holland, 2021; The Economist, 2021).

Earlier studies on P2P lending have raised concerns about future interest rate increases. Tomlinson et al. (2016) estimated that the interest rate environment may lead to an expected divergence in the penetration of P2P lending in the UK market from £0.5 billion (under a normalised interest rate environment) to £35.5 billion (under the current interest rate environment) by 2025. Our study

Table 10
Impact of stock market returns and interest rates on probability of default: Role of aggregate loan volumes and euro area interest rates.

Panel A: Subsamples based on median loan volume				
	Lower than median loan volume	Higher than median loan volume	Lower than median loan volume	Higher than median loan volume
Variables	DV=*BADLOAN*	DV=*BADLOAN*	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET	0.0643**	0.0690**		
	(0.0297)	(0.0297)		
INTRATE			0.4306*	0.1705
			(0.2491)	(0.2410)
Controls	Yes	Yes	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes	Yes	Yes
R-squared	0.1755	0.1890	0.1726	0.1897
N	2,135,510	2,112,951	2,135,510	2,112,951
Two-sample coefficient test: Chow test chi-squared statistic				
MARKETRET/ AVEINTRATE	0.71		0.68	

Panel B: Subsamples based on annualised agreed rate (*AAR*)				
	Higher than median *AAR*	Lower than median *AAR*	Higher than median *AAR*	Lower than median *AAR*
Variables	DV=*BADLOAN*	DV=*BADLOAN*	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET	0.3301***	0.0872***		
	(0.0188)	(0.0121)		
INTRATE			0.8346***	0.7447***
			(0.0052)	(0.0069)
Controls	Yes	Yes	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes	Yes	Yes
R-squared	0.0158	0.0242	0.0796	0.1133
N	2,642,170	2,485,496	1,616,247	1,224,297
Two-sample coefficient test: Chow test chi-squared statistic				
AVEINTRATE/ MARKETRET	63.34***		52.48**	

Note: Table 10 presents the baseline regression results for two panels. Panel A draws subsamples based on the criteria of loan volumes in countries being higher or lower than the median level. Panel B draws subsamples based on the annualised agreed rate (*AAR*) by credit and other institutions. All model specifications employ robust standard errors in parentheses; $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

highlights that ECB interest rate increases may significantly increase the sensitivity of P2P loan defaults to financial market volatility.

Another specific aspect of our sample is that it incorporates countries that are diverse in terms of their credit market's efficiency and the progress of their FinTech development. A unique opportunity is created by the similarity of these countries due to their geographical location and the fact that they operate under the EU jurisdiction, allowing exploration of the determinants of defaults in the breakdown of financial markets, as well as FinTech development. Therefore, we analyse the data set in panels consisting of subsamples, based on FinTech adoption and NPLs in the banking sector.

Table 11, Panel A provides the breakdown of our baseline model into panels based on the level of FinTech adoption in individual countries. We divide our database based on the Global FinTech Index, as reported by Findexable (2019), which provides a snapshot of local business infrastructure and FinTech ecosystem quality. The impacts of interest rates and market returns are significantly different between countries with higher than median level Global FinTech Index scores and those with scores lower than the median level. Thus, the transfer of risk from conventional financial markets to the P2P lending market is more pronounced when the country has a high level of FinTech adoption.

Table 11, Panel B presents another perspective on the interconnectedness between the P2P lending market and conventional financial markets by providing an analysis of the breakdown of NPLs in the banking sector. We divide our database into two non-performing loan (NPL) subsamples, those higher than the median level and those lower than the median level. We observe that under high levels of NPLs, P2P loans are relatively more sensitive to changes in market returns and interest rates.

In addition to subsample analysis, we analyse the relationship between traditional financial markets and FinTech loans using interaction variables, creating interactions of euro area interest rates, banking sector NPLs and FinTech adoption with our two variables of interest. Table 12 presents the regression results based on these three interaction variables. The euro area benchmark interest rate and banking sector NPLs are found to significantly amplify the impact of both market returns and interest rates on the probability of default. In contrast, FinTech adoption has a mixed impact on FinTech loans. Specifically, we observe that the interaction between FinTech adoption and interest rates has a negative impact on the probability of default among FinTech loans. These findings shed more light on the ongoing debate on the relationship between traditional and FinTech lending markets (Cole et al., 2019; de Roure et al., 2022), by indicating that FinTech lending complements traditional financial lending.

The issue of the relationship between traditional and P2P lending markets requires consideration of the possibility of borrower self-selection. Do higher interest rates or higher market returns cause higher loan delinquency or are the borrowers who take a higher level

Table 11
Impact of stock market returns and interest rates on probability of default: Role of FinTech adoption and banking sector non-performing loans (NPLs).

Panel A: Subsamples based on FinTech adoption				
	High level of FinTech adoption	High level of FinTech adoption	High level of FinTech adoption	High level of FinTech adoption
Variables	DV=*BADLOAN*	DV=*BADLOAN*	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET	0.4672***	0.1118***		
	(0.0156)	(0.0132)		
INTRATE			0.4138***	0.3262***
			(0.0063)	(0.0262)
Controls	Yes	Yes	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes	Yes	Yes
R-squared	0.0182	0.0263	0.0423	0.0262
N	1,791,040	2,568,524	1,791,040	2,568,524
Two-sample coefficient test: Chow test chi-squared statistic				
MARKETRET/INTRATE	80.29***		11.81***	

Panel B: Subsamples based on banking sector NPLs				
	High level of NPLs	High level of NPLs	High level of NPLs	High level of NPLs
Variables	DV=*BADLOAN*	DV=*BADLOAN*	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET	0.7582***	0.0968***		
	(0.0226)	(0.0118)		
INTRATE			0.7674***	0.7010***
			(0.0082)	(0.0059)
Controls	Yes	Yes	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes	Yes	Yes
R-squared	0.0184	0.0383	0.0862	0.0961
N	2,103,877	2,240,850	2,103,877	2,240,850
Two-sample coefficient test: Chow test chi-squared statistic				
MARKETRET/INTRATE	146.23***		63.63***	

Note: Table 11 presents the baseline regression results for two panels. Panel A reports the findings of logit regression analysis for countries with high and low levels of FinTech adoption. The subsamples are based on countries' FinTech Development Index (Findexable, 2019) being higher/lower than the global median. Panel B draws subsamples based on gross non-performing loans (NPLs), dividing the sample into high and low levels of NPLs (compared to the median) in the banking sector. All model specifications employ robust standard errors in parentheses: $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

of risk simply attracted to borrow from that platform? The causation perspective in an empirical setting cannot disentangle the two effects. One can argue that the shift in the borrower's creditworthiness is the result of an increased (or decreased) loan rating mechanically causing changes in a default or overdue loan. Additionally, the platform may take an active role in effectively managing the loan to fine tune the loan quality.

Accordingly, Table 13 provides the results of the analysis using subsamples based on loan ratings. We divide the database into three groups: Panel A ('A' and 'A-' rating loans), Panel B ('B+' and 'B-' rating loans) and Panel C ('B-', 'C+' and 'D' rating loans). All the regression models reported in Table 13 generate significant positive coefficients for both market returns and interest rates. Results from the breakdown of borrower loan ratings also indicate significantly high sensitivity of loans in Panel C, compared to those in Panels A and B, to changes in market returns and interest rates.

To assess the change in default risk based on loan ratings, we estimate the marginal magnitudes of changes from the effect of the independent variables for the three rating subgroups. We hold all the control variables at their mean values and estimate the change in the likelihood of default under the respective changes in market returns and interest rates. We base our parameters on the findings reported in Table 13. By estimating the change in the likelihood of default for the three rating subgroups, we find that the likelihood of default is greater for the 'B-, C+ and D' rating subgroup for each marginal change in market returns. The distinguishing feature of marginal effects, as reported in Fig. 1, is the similar sensitivity of 'A and 'B' rating loans to market returns, with 'B+ and B' rating loans being slightly less sensitive. Fig. 1 also reports the respective marginal changes in the probability of default due to interest rates in the breakdown of borrower ratings. We observe that 'B-, C+ and D' rating loans are the most sensitive category in terms of interest rates, compared to the other borrower rating categories. On average, loans in the lowest rating category are 0.80 % more likely to default when market returns increase and 0.37 % more likely to default in the case of an increase in interest rates (compared to other rating categories).

Another important issue in credit risk assessment is the time of loan closure, the early occurrence of which has a substantial cost for investors. Our study examines this specific aspect of loan defaults using survival analysis by estimating not only whether, but also when, the default event occurs (Royston & Lambert, 2011). The model, based on survival analysis, uses the same set of country-specific variables together with a wide range of borrower-specific variables. These authors used, as their analysis's dependent variable, the time to the occurrence of an event of interest, that is, a default. In the current study, the dependent variable is the risk of failure or how long the loan survived until the failure (default). This is done by means of the Cox proportional hazards regression which relates survival

Table 12
Impact of stock market returns and interest rates on probability of default: Interaction effects.

Panel A: Interaction variables with euro area interest rate		
Variables	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET × *ECB_RATE*	0.6702***	
	(0.0054)	
INTRATE × *ECB_RATE*		1.7711***
		(0.0090)
Controls	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes
R-squared	0.0220	0.0899
N	4,485,532	4,485,532
Panel B: Interaction variables with banking sector NPLs		
Variables	**DV=*BADLOAN***	**DV=*BADLOAN***
MARKETRET × *NPL*	0.0184***	
	(0.0002)	
INTRATE × *NPL*		0.1096***
		(0.0013)
Controls	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes
R-squared	0.0177	0.0659
N	4,485,532	4,485,532
Panel C: Interaction variables with FinTech adoption		
Variables	**DV=*BADLOAN***	**DV=*BADLOAN***
MARKETRET × *FINTECH*	0.0068***	
	(0.0000)	
INTRATE × *FINTECH*		−0.0049***
		(0.0002)
Controls	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes
R-squared	0.0278	0.0700
N	4,485,532	4,485,532

Note: Table 12 presents the baseline regression results for three panels. Panel A reports the findings of logit regression analysis for the interaction variable with the euro area interest rate. Panel B reports the findings for the interaction variable with the banking sector NPLs. Panel C reports the findings for the interaction variable with the FinTech Development Index (Findexable, 2019). All model specifications employ robust standard errors in parentheses: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

time and explanatory variables.

Table 14, Panel A provides the survival analysis results, by means of Cox regressions, one for each explanatory variable. The regression coefficient is interpreted as a k-fold increase in risk, with a positive regression coefficient meaning that the risk is higher. The risk ratio can be interpreted as the predicted change in the risk for a unit increase in the explanatory variable. Table 14 reveals important practical findings. Both variables of interest are consistent across the models. Model (1) estimates that an increase of one unit in market returns increases the risk of loan failure by 0.0476 times, ceteris paribus. In contrast, the risk of failure increases by 0.3617 times with each percentage increase in interest rates. We further investigate the survival of loans by dividing them into two subsamples based on the time of loan closure. Table 14, Panel B indicates that the probability of failure is significantly higher for loans during the first 100 days from the issue date.

Survival curves are another important aspect of survival analysis, with these indicating the probability of failure at a certain point of time. Fig. 2 displays the survival curves for market returns as well as for interest rates. The probability of loan survival is almost identical during the first 100 days from the issue of the loan when a decrease or increase occurs in market returns. However, a significant drop in survivability is observed immediately following the 100-day mark when an increase occurs in market returns. We also observe that loan survival is more likely during a low interest period compared to a high interest period ('high' and 'low' are with respect to the median level).

5. Concluding remarks

5.1. Theoretical and practical implications

This study investigates the effect of various determinants on P2P loan delinquencies in a cross-country study. Unlike the limited number of earlier studies (Li et al., 2020; Serrano-Cinca et al., 2015; Wei & Lin, 2016), we rely on an extended cross-country data set with a wide range of country-specific variables. This extension allows the findings to be generalised to a broader international context. The findings also help investors and P2P platforms to better estimate country risk in their risk assessments. Thus, several theoretical and practical implications of this study can be highlighted.

Table 13
Impact of stock market returns and interest rates on probability of default: Rating subsamples.

Panel A: 'A' & 'A-' rating loans		
	(1)	(2)
Variables	DV=*BADLOAN*	DV=*BADLOAN*
MARKETRET	0.7872***	
	(0.0905)	
INTRATE		1.5521***
		(0.0333)
Controls	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes
R-squared	0.0433	0.0512
N	945,988	945,988
Panel B: 'B+' & 'B' rating loans		
	(1)	**(2)**
Variables	**DV=*BADLOAN***	**DV=*BADLOAN***
MARKETRET	0.2401***	
	(0.0392)	
INTRATE		1.4546***
		(0.0442)
Controls	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes
R-squared	0.0602	0.1073
N	1,282,510	1,282,510
Panel C: 'B-', 'C+' and 'D' rating loans		
	(1)	**(2)**
Variables	**DV=*BADLOAN***	**DV=*BADLOAN***
MARKETRET	3.8771***	
	(0.151)	
INTRATE		3.5432***
		(0.2185)
Controls	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes
R-squared	0.0572	0.0464
N	1,629,720	1,629,720
Test of equality of coefficients: Chow test chi-squared statistic		
MARKETRET/INTRATE	74.34***	86.67***

Note: Table 13 presents the results of regression analyses based on three panels (by loan ratings). These are the logit regression analysis results for the likelihood of loan default (*BADLOAN*). All model specifications employ robust standard errors in parentheses; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Firstly, a key contribution of this study is its cross-country model that combines various econometric and machine learning methods with a multi-million observation data set. Specifically, we applied the LASSO method for model selection, parameter estimation and instrumental variable selection. By integrating loan-specific and economy-specific variables, the model provides a robust risk management framework. The findings reveal specific odds ratios for market returns and interest rates, enabling managers to assess credit risk more accurately and to effectively adjust their strategies.

Secondly, our findings reveal that borrowers' creditworthiness in the P2P lending market is closely linked to broader financial markets. This insight allows platform managers to better manage and mitigate various risk factors contributing to borrower default risks. By understanding the sensitivity of default risk to external factors, such as stock market returns and interest rates, managers can implement proactive risk mitigation strategies, thereby enhancing platform stability and boosting investor confidence.

Thirdly, a key theoretical contribution of our study is its exploration of the 'financial accelerator' effect and life-cycle consumption theory. The study demonstrates that increased borrowing levels elevate consumption rates and borrowers' debts. Researchers can investigate behavioural patterns related to credit risk sensitivity using the coefficients of variables, stock market return (*MARKETRET*) and interest rate (*INTRATE*). These coefficients are valuable for stress testing P2P lending portfolios, enabling platforms to simulate scenarios (e.g., increased stock market volatility), assess their resilience and develop effective risk mitigation strategies.

Fourthly, the findings of our study make an important theoretical contribution to the debate on the relationship between traditional and FinTech lending markets, as well as the risk of transfer between them. Specifically, our study empirically indicated that FinTech lending complements traditional financial lending. We have also provided empirical evidence that stock market conditions could bilaterally influence default risk in the P2P lending market.

Finally, the findings about country-level indicators may have significant implications for better understanding the interaction between P2P lending and the broader economy. In doing so, the study's findings contribute to risk theory by emphasising the interconnectedness of financial markets and P2P lending. Incorporating both loan-specific and broader economic variables enriches

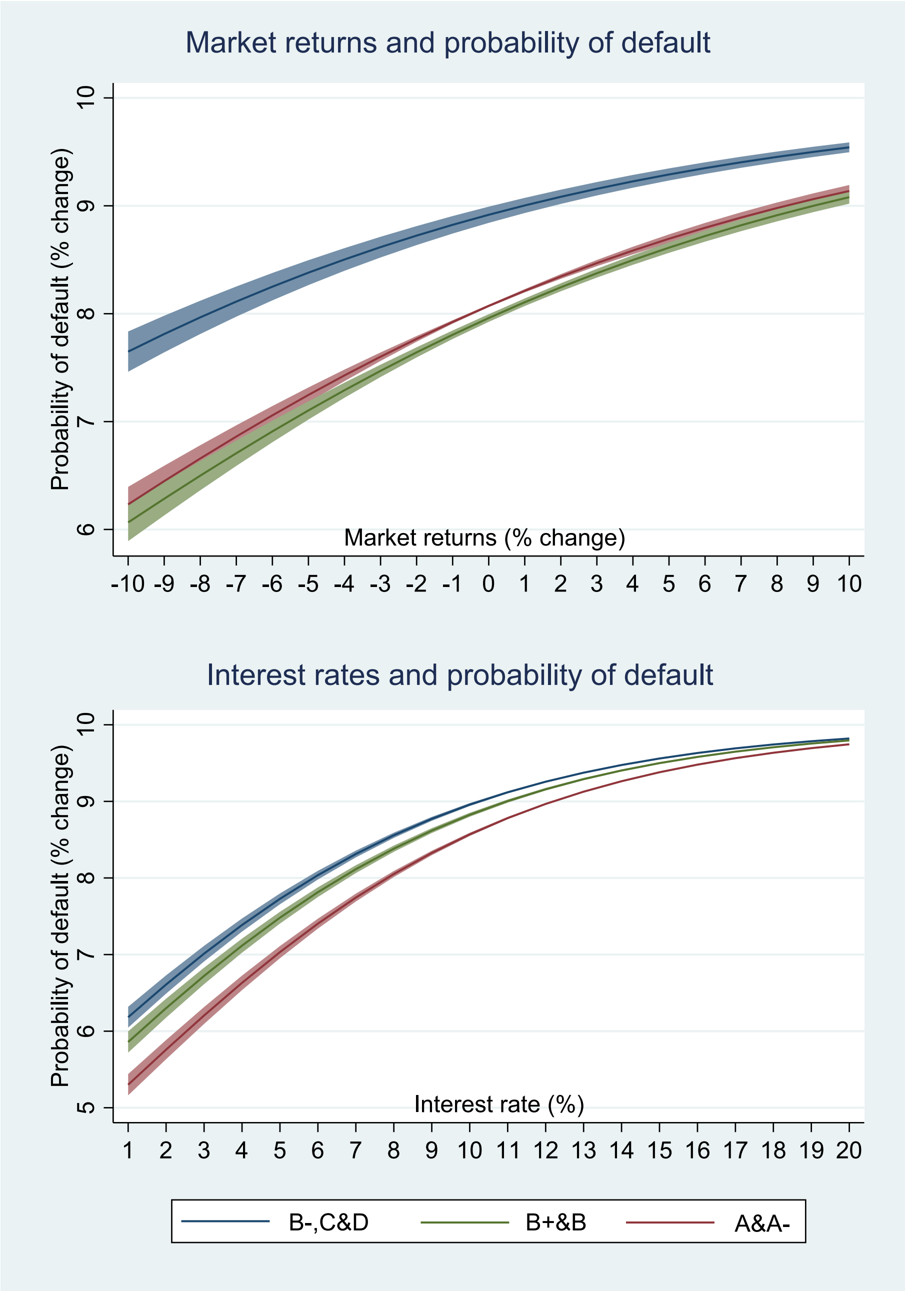

Fig. 1. Change in probability of default for incremental change in explanatory variables (by loan ratings).
Note: Fig. 1 presents estimates of the change in the likelihood of default for each incremental increase in explanatory variables by holding all control variables at their mean values. Estimation parameters are based on the findings reported in Table 13 in the breakdown of borrower rating groups. The shaded area represents 95 % confidence intervals for the respective estimations.

existing risk frameworks. Researchers can build upon this foundation, exploring additional risk determinants and refining theories applicable to the FinTech sector.

5.2. Generalisation of research findings, limitations and room for further research

Using the cross-country database, we find that country-level interest rate and stock market returns significantly impact default risk. An increase in interest rate and stock market returns leads to higher aggregate default rates at the country level. Our study's main findings are consistent across robustness tests based on endogeneity correction, LASSO models and adjustments for selection bias. We

Table 14
Impact of stock market returns and interest rates on probability of default: Survival time of issued loans.

Panel A: Survival analysis				
	(1)		(2)	
	DV=*SURVIVAL_TIME*		DV=*SURVIVAL_TIME*	
	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio
MARKETRET	3.0453***	0.0476***		
	(0.0128)	(0.0006)		
INTRATE			1.2857***	0.36171***
			(0.0023)	(0.0084)
LOANVOL	0.8609***	−0.1498***	0.8555***	−0.1561***
	(0.0005)	(0.0006)	(0.0005)	(0.0006)
AVEAMNT	0.9931***	−1.0501***	1.2209***	−1.0648***
	(0.0001)	(0.0010)	(0.0036)	(0.0010)
AVEDUR	0.3499***	0.0351***	0.3448***	0.0654***
	(0.0003)	(0.0009)	(0.0003)	(0.0009)
AVERATING	1.5170***	−0.0731***	1.2209***	−0.3728***
	(0.0044)	(0.0041)	(0.0036)	(0.0041)
EARNINGS	0.9295***	0.8412***	0.6888***	0.8692***
	(0.0038)	(0.0012)	(0.0028)	(0.0012)
GDP	2.3190***	5.0224***	2.3851***	8.9978***
	(0.0028)	(0.0307)	(0.0028)	(0.0288)
ESI	1.7745***	−0.0369***	8.3947***	−0.0658***
	(0.0058)	(0.0003)	(0.0051)	(0.0003)
INFLATION	0.9638***	0.0921***	0.9363***	0.0757***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
NPL	1.0965***	−0.0420***	1.0786***	0.0085***
	(0.0002)	(0.0004)	(0.0002)	(0.0004)
Pseudo-*R*-squared	0.0348		0.0372	
N	3,829,464		3,829,464	

Panel B: Subsamples based on loan closure				
	(1)		**(2)**	
	DV=*BADLOAN*		**DV=*BADLOAN***	
Variables	**Up to 100 days**	**More than 100 days**	**Up to 100 days**	**More than 100 days**
MARKETRET	0.9516***	0.3993***		
	(0.0884)	(0.0355)		
INTRATE			1.7679***	0.3552***
			(0.0555)	(0.0243)
Controls	Yes	Yes	Yes	Yes
Constant, Yr. & Ind. effects	Yes	Yes	Yes	Yes
R-squared	0.0544	0.0585	0.0913	0.0607
N	3,484,956	475,474	3,484,956	475,474
Test of equality of coefficients: Chow test chi-squared statistic				
MARKETRET/INTRATE	22.09***		11.40***	

Note: Table 14 presents the results from the survival time of issued loans. Panel A provides the Cox regression analysis for the survival time of issued loans. Panel B provides the results for subsamples based on realised loan closure. Refer to Table 1 for descriptions of variables. Number of listings analysed: 3,829,464. Failed: 736,406 (19.23 %). Successful: 3,093,058 (80.77 %). All model specifications employ robust standard errors in parentheses; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

also find no significant differences in the functional form of interest rates between subsamples, based on the ECB base interest rates. However, we observe that the impact of stock market returns is significantly higher for non-zero interest rate periods compared to zero interest rate periods (based on the ECB base interest rates). The severity of the impact of market returns and interest rates is found to be significantly different based on the level of FinTech adoption as well as the level of banking sector distress.

Our study makes the following remarks on the generalisability of its findings and its limitations. Although the Mintos platform used in this study is based in Latvia, it issues loans and accepts investments across the European Union (EU). Therefore, the findings should be relevant to other P2P lending platforms in the European Union (EU). The time span of the database in our study extends to five years which could be problematic in capturing macroeconomic tendencies. However, the model's framework is developed from well-proven traditional financial markets, with investor incentives not generally changing. Thus, the generalisation of findings to 2020 and beyond should also be relevant.

Several aspects with regard to macroeconomic indicators, such as GDP growth, and their impact on default probability require this study to make several reservations. Firstly, existing studies have provided rather contradictory results on the impact of macroeconomic factors on defaults. For instance, Croux et al. (2020) and Louzis et al. (2012) indicated a negative relationship of GDP growth rate with consumer and mortgage NPLs, as well as with the delinquency of P2P loans. On the other hand, the study of Avgeri and Psillaki (2023) indicated a positive relationship with loan defaults, implying that an improvement in the growth of GDP results in a higher likelihood

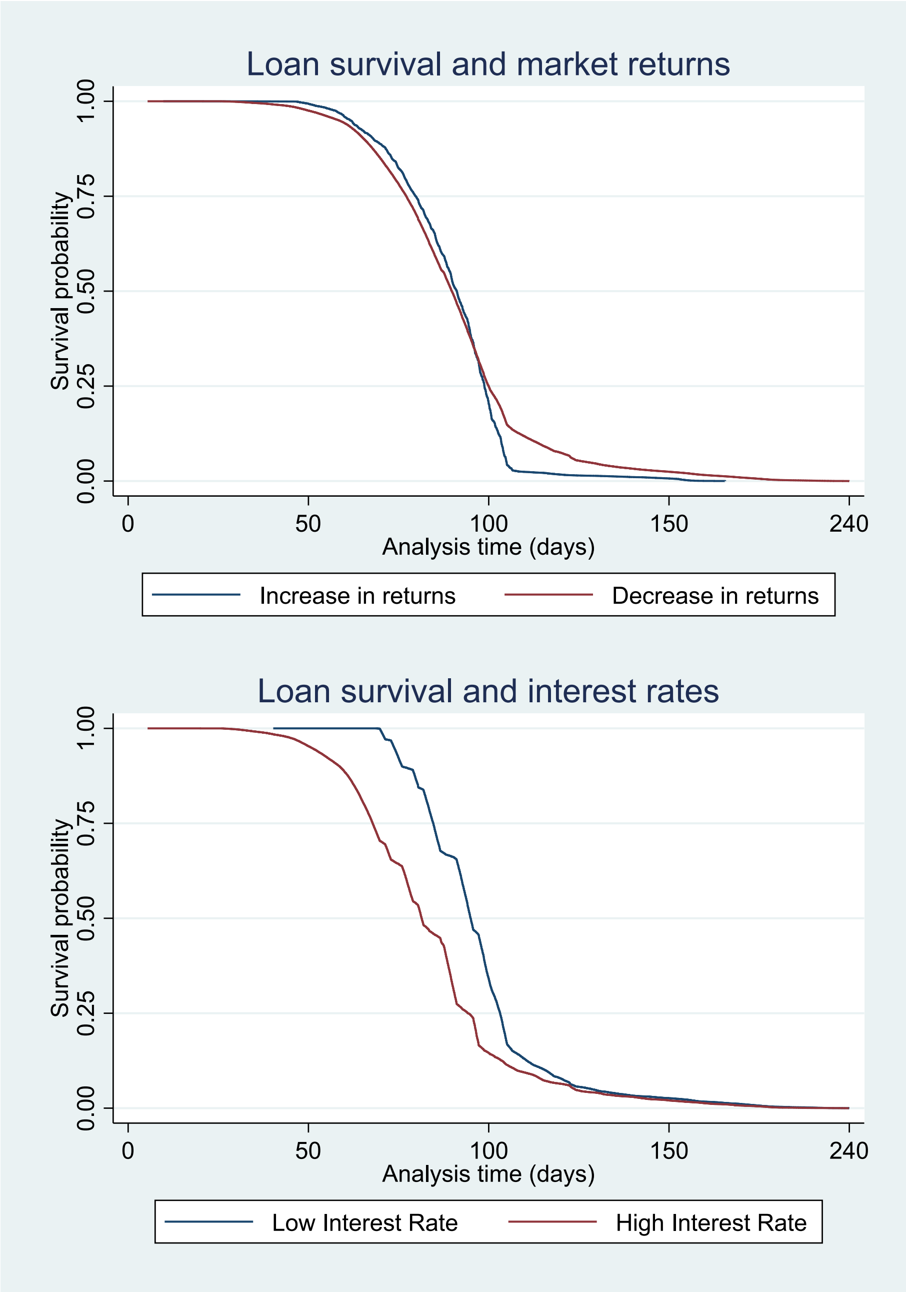

Fig. 2. Survival function for respective groups based on explanatory variables.
Note: Fig. 2 presents estimates of the survival function for subsamples based on interest rates and market returns. Estimation parameters are based on the findings reported in Table 14.

of delinquency. This could be attributed to the trade-off between economic growth and increasing debt in the economy.

Secondly, if macroeconomic conditions change, such as the GFC, defaults are no longer independent events and require a different framework for analysis. The impact of macroeconomic factors, such as GDP growth, could be better analysed under conditions of economic distress, as in the study of Nigmonov and Shams (2021). Finally, the diversity of countries in the database used in this study in terms of geographical scope and FinTech development requires further evaluation. Existing studies that provided consistent evidence on the impact of macroeconomic variables on defaults (e.g., Figlewski et al., 2012; Kukuk & Rönnberg, 2013) considered corporate defaults in a single country and/or in traditional financial markets. However, due to limited available data, it was not possible to further investigate the control variables. Although this issue was beyond the scope of the current study, it may provide some direction towards promising fields for further research.

The study's findings should also be applied with caution during periods of financial distress. The P2P lending market experienced a challenging period during 2020. The COVID-19 pandemic and related turmoil in traditional financial markets, as well the resulting economic recession, created significant problems for the burgeoning P2P lending market. This study mainly explores the insolvency of borrowers which is expected to skyrocket as the current adverse economic conditions persist. The study's findings on macroeconomic variables may offer some clues about the expected impact of economic conditions on the P2P lending market. At the same time, we cannot observe the full scale of the problems related to insolvency in the short term, as various government policies are shielding borrowers and providing liquidity to investors. Additionally, the database used in this study proved to have a limited scope when analysing subsamples.

Considering the limitations of this research, future studies could enhance and extend this study in several directions. Firstly, the LASSO machine learning method could be enhanced with further extensions. Machine learning models, such as XGBoost and neural networks, tend to surpass traditional ones, such as logistic regression and linear discriminant analysis. However, the causal interpretation of machine learning model parameters remains a challenge for future research. Secondly, understanding how machine learning models perform over time is of interest, especially during financial crises when default rates rapidly increase, such as the liquidity risk faced by platforms and investors in the short term. Therefore, future studies could explore the liquidity risk in the P2P lending market when faced with contagion-type external conditions. Finally, future studies could extend the database by collecting data from a larger number of platforms. In doing so, these studies could investigate tendencies that were not explored in this study.

CRediT authorship contribution statement

Asror Nigmonov: Conceptualization, Methodology, Data curation, Writing – original draft, Visualization, Formal analysis, Software. **Syed Shams:** Supervision, Validation, Writing – review & editing. **Povilas Urbonas:** Data curation, Software, Writing – review & editing.

Declaration of competing interest

None.

Data availability

Data will be made available on request.

Appendix A. Appendix

Table A1
Breakdown of loans by country and loan originators.

Country	Loan originator	N	%
Bulgaria	CashCredit	76,263	1.70 %
Bulgaria	Credissimo	26,912	0.60 %
Bulgaria	ITF Group	62,978	1.40 %
Bulgaria	Mogo	9267	0.21 %
Bulgaria	StikCredit	4070	0.09 %
Denmark	Creamfinance	48,692	1.09 %
Denmark	Mozipo Group	2080	0.05 %
Denmark	Simbo	282,494	6.30 %
Estonia	Capitalia	252	0.01 %
Estonia	Creditstar	29,677	0.66 %
Estonia	ESTO	14,644	0.33 %
Estonia	Mogo	22,393	0.50 %
Estonia	Placet	17,387	0.39 %
Finland	BB Finance Group	42,431	0.95 %
Finland	Creditstar	11,074	0.25 %
Latvia	AgroCredit	202	0.00 %
Latvia	Banknote	854,288	19.05 %
Latvia	Bino	692,587	15.44 %
Latvia	Capitalia	1741	0.04 %
Latvia	Creamfinance	17,086	0.38 %
Latvia	Hipocredit	1255	0.03 %
Latvia	Mogo	62,385	1.39 %
Latvia	Mogo Renti	7032	0.16 %
Latvia	VIZIA	25,660	0.57 %
Lithuania	Capitalia	1827	0.04 %

(*continued on next page*)

Table A1 (*continued*)

Country	Loan originator	N	%
Lithuania	Debifo	2967	0.07 %
Lithuania	EBV Finance	41,244	0.92 %
Lithuania	Hipocredit	229	0.01 %
Lithuania	Mogo	29,801	0.66 %
Lithuania	Mozipo Group	51,899	1.16 %
Lithuania	Placet	18,637	0.42 %
Poland	Aasa	35,196	0.78 %
Poland	Aforti	1049	0.02 %
Poland	Alfakredyt	30,936	0.69 %
Poland	Capital Service	105,919	2.36 %
Poland	Creamfinance Poland	34,973	0.78 %
Poland	Creditstar	152,298	3.40 %
Poland	Dziesiatka Finanse	3593	0.08 %
Poland	Efaktor	100	0.00 %
Poland	Eurocent	1392	0.03 %
Poland	Everest Finanse	89,936	2.01 %
Poland	GetBucks	11,129	0.25 %
Poland	Invipay	932	0.02 %
Poland	Kuki	644,027	14.36 %
Poland	Leaselink	1953	0.04 %
Poland	Mogo	5054	0.11 %
Romania	Credius	25,303	0.56 %
Romania	Extra Finance	79	0.00 %
Romania	Mikro Kapital Romania	260	0.01 %
Romania	Mogo	7206	0.16 %
Romania	Mozipo Group	10,344	0.23 %
Spain	Creamfinance Spain	19,318	0.43 %
Spain	Creditstar	66,073	1.47 %
Spain	Dindin	12,878	0.29 %
Spain	Dineo Credito	60,196	1.34 %
Spain	Fireof	63	0.00 %
Spain	ID Finance	452,319	10.08 %
Spain	Lendrock	1	0.00 %
Spain	Rapido Finance	82,109	1.83 %
Sweden	Aasa	6831	0.15 %
United Kingdom	1 pm	19	0.00 %
United Kingdom	Evergreen	3655	0.08 %
United Kingdom	Novaloans	21,414	0.48 %
United Kingdom	Peachy	139,523	3.11 %
Total		**4,485,532**	**100.00 %**

Table A2
Breakdown of loans by rating and loan status.

Rating	N	%
A	143,157	3.19 %
A-	1,100,013	24.52 %
B	1,008,766	22.49 %
B+	108,285	2.41 %
B-	1,748,919	38.99 %
C	13,889	0.31 %
C+	127,301	2.84 %
C-	11,129	0.25 %
D	224,073	5.00 %
Total	**4,485,532**	**100.00 %**
Loan status	**N**	**%**
Current	93,060	2.07 %
Default	85	0.00 %
Finished - Bad Debt	152	0.00 %
Finished as scheduled	442,645	9.87 %
Finished prematurely	3,361,770	74.95 %
Finished prematurely - Buyback	555,418	12.38 %
Grace Period	2206	0.05 %
Late 1–15	9569	0.21 %
Late 16–30	9089	0.20 %
Late 31–60	9373	0.21 %
Late 60+	2165	0.05 %
Total	**4,485,532**	**100.00 %**

References

Akerlof, G. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics, 84*(3), 488–500.

Akhtaruzzaman, M., Boubaker, S., & Sensoy, A. (2021). Financial contagion during COVID-19 crisis. *Finance Research Letters, 38*, Article 101604.

Ali, A., & Daly, K. (2010). Macroeconomic determinants of credit risk: Recent evidence from a cross country study. *International Review of Financial Analysis, 19*(3), 165–171.

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

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609.

Amemiya, T. (1978). The estimation of a simultaneous equation generalized probit model. *Econometrica: Journal of the Econometric Society, 46*, 1193–1205.

Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics, 11*, 727–753.

Aretz, K., Florackis, C., & Kostakis, A. (2018). Do stock returns really decrease with default risk? New international evidence. *Management Science, 64*(8), 3821–3842.

Avgeri, E., & Psillaki, M. (2023). Factors determining default in P2P lending. *Journal of Economics Studies, 51*(4), 823–840.

Babaei, G., Giudici, P., & Raffinetti, E. (2023). Explainable FinTech lending. *Journal of Economics and Business, 125*, Article 106126.

Balyuk, T., & Davydenko, S. (2019). Reintermediation in FinTech: Evidence from online lending. *Journal of Financial and Quantitative Analysis*, 1–75.

Bebczuk, R. N. (2003). *Asymmetric information in financial markets: Introduction and applications*. Cambridge University Press.

Beck, T. (2013). Bank financing for SMEs – Lessons from the literature. *National Institute Economic Review, 225*(1), R23–R38.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014a). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives, 28*(2), 29–50.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014b). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies, 81*(2), 608–650.

Belloni, A., Chernozhukov, V., & Wei, Y. (2016). Post-selection inference for generalized linear models with many controls. *Journal of Business & Economic Statistics, 34*(4), 606–619.

Berg, T., Fuster, A., & Puri, M. (2022). FinTech lending. *Annual Review of Financial Economics, 14*, 187–207.

Bergmann, M. (2020). The determinants of mortgage defaults in Australia: Evidence for the double-trigger hypothesis. In *Research discussion paper (RDP) 2020–03*. Reserve Bank of Australia.

Bernanke, B., & Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *The American Economic Review, 79*(1), 14–31.

Bernanke, B., Gertler, M., & Gilchrist, S. (1998). *The financial accelerator in a quantitative business cycle framework*. National Bureau of Economic Research (NBER) Working Paper Series (No. 6455).

Bester, H. (1985). Screening vs. rationing in credit markets with imperfect information. *The American Economic Review, 75*(4), 850–855.

Bocola, L., Bornstein, G., & Dovis, A. (2019). Quantitative sovereign default models and the European debt crisis. *Journal of International Economics, 118*, 20–30.

Bonfim, D. (2009). Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of Banking & Finance, 33*(2), 281–299.

Boot, A. W., & Thakor, A. V. (2000). Can relationship banking survive competition? *The Journal of Finance, 55*(2), 679–713.

Boubaker, S., Farag, H., & Nguyen, D. K. (2015). Short-term overreaction to specific events: Evidence from an emerging market. *Research in International Business and Finance, 35*, 153–165.

Bu, Y., Du, X., Li, H., Yu, X., & Wang, Y. (2023). Research on the FinTech risk early warning based on the MS-VAR model: An empirical analysis in China. *Global Finance Journal, 58*, Article 100898.

Buchak, G., Matvos, G., Piskorski, T., & Seru, A. (2018). FinTech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics, 130*(3), 453–483.

Bühlmann, P., & van de Geer, S. (2011). *Statistics for high-dimensional data: Methods, theory and applications*. Springer Science & Business Media.

Cai, S., Lin, X., Xu, D., & Fu, X. (2016). Judging online peer-to-peer lending behavior: A comparison of first-time and repeated borrowing requests. *Information & Management, 53*(7), 857–867.

Cambridge Centre for Alternative Finance (CCAF). (2024). *The future of global FinTech: Towards resilient and inclusive growth. 18 January*. Insight Report. World Economic Forum.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance, 63*(6), 2899–2939.

Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of Finance, 8*(4), 537–569.

Chen, M. Y. (2011). Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. *Computers & Mathematics with Applications, 62*(12), 4514–4524.

Chen, P., & Wu, C. (2014). Default prediction with dynamic sectoral and macroeconomic frailties. *Journal of Banking & Finance, 40*, 211–226.

Chernick, M. R., & LaBudde, R. A. (2014). *An introduction to bootstrap methods with applications to R*. John Wiley & Sons.

Chernozhukov, V., Hansen, C. B., & Spindler, M. (2015). Valid post-selection and post-regularization inference: An elementary, general approach. *Annual Review of Economics, 7*, 649–688. https://doi.org/10.1146/annurev-economics012315-015826

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society, 28*(3), 591–605.

Cole, R., Cumming, D., & Taylor, J. (2019). *Does FinTech compete with or complement bank finance?*. Available at SSRN 3302975.

Collin-Dufresne, P., & Goldstein, R. S. (2001). Do credit spreads reflect stationary leverage ratios? *The Journal of Finance, 56*(5), 1929–1957.

Cornett, M. M., McNutt, J. J., Strahan, P. E., & Tehranian, H. (2011). Liquidity risk management and credit supply in the financial crisis. *Journal of Financial Economics, 101*(2), 297–312.

Croux, C., Jagtiani, J., Korivi, T., & Vulanovic, M. (2020). Important factors determining FinTech loan default: Evidence from a LendingClub consumer platform. *Journal of Economic Behavior and Organization, 173*, 270–296.

Degryse, H., & Ongena, S. (2007). The impact of competition on bank orientation. *Journal of Financial Intermediation, 16*(3), 399–424.

Dinger, V. (2009). Do foreign-owned banks affect banking system liquidity risk? *Journal of Comparative Economics, 37*(4), 647–657.

Dorfleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., de Castro, I., & Kammler, J. (2016). Description-text related soft information in peer-to-peer lending – Evidence from two leading European platforms. *Journal of Banking & Finance, 64*, 169–187.

Duan, J. C., Sun, J., & Wang, T. (2012). Multiperiod corporate default prediction – A forward intensity approach. *Journal of Econometrics, 170*(1), 191–209.

Duffie, D., Saita, L., & Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics, 83*(3), 635–665.

Dushnitsky, G., Guerini, M., Piva, E., & Rossi-Lamastra, C. (2016). Crowdfunding in Europe: Determinants of platform creation across countries. *California Management Review, 58*(2), 44–71.

Dyreng, S. D., Mayew, W. J., & Williams, C. D. (2012). Religious social norms and corporate financial reporting. *Journal of Business Finance & Accounting, 39*(7–8), 845–875.

Eça, A., Ferreira, M. A., Prado, M. P., & Rizzo, A. E. (2021). The real effects of FinTech lending on SMEs: Evidence from loan applications. In *CEPR [Centre for Economic Policy Research] Discussion Paper No. DP16684*. Available at SSRN: https://ssrn.com/abstract=3960288.

Emekter, R., Tu, Y., Jirasakuldech, B., & Lu, M. (2015). Evaluating credit risk and loan performance in online peer-to-peer (P2P) lending. *Applied Economics, 47*(1), 54–70.

Ericsson, J., Jacobs, K., & Oviedo, R. (2009). The determinants of credit default swap premia. *Journal of Financial and Quantitative Analysis, 44*(1), 109–132.

Espinoza, R. A., & Prasad, A. (2010). *Non-performing loans in the GCC banking system and their macroeconomic effects*. International Monetary Fund.

European Central Bank (ECB). (2021). Key ECB interest rates from European Central Bank. https://www.ecb.europa.eu/stats/policy_and_exchange_rates/key_ecb_interest_rates/html/index.en.html.

Félix, E. G. S., Pires, C. P., & Gulamhussen, M. A. (2013). The determinants of venture capital in Europe – Evidence across countries. *Journal of Financial Services Research, 44*(3), 259–279.

Figlewski, S., Frydman, H., & Liang, W. (2012). Modeling the effect of macroeconomic factors on corporate default and credit rating transitions. *International Review of Economics and Finance, 21*(1), 87–105.
Findexable. (2019). The global Fintech index 2020. Findexable Limited & the Global Fintech Index. Retrieved from: www.findexable.com.
Freue, G. V. C., Ortiz-Molina, H., & Zamar, R. H. (2013). A natural robustification of the ordinary instrumental variables estimator. *Biometrics, 69*(3), 641–650.
Fuster, A., & Willen, P. S. (2017). Payment size, negative equity, and mortgage default. *American Economic Journal: Economic Policy, 9*(4), 167–191.
Gai, P., & Kapadia, S. (2010). Contagion in financial networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 466*(2120), 2401–2423.
Galema, R. (2020). Credit rationing in P2P lending to SMEs: Do lender–borrower relationships matter? *Journal of Corporate Finance, 65*, Article 101742.
Ghosh, A. (2015). Banking-industry specific and regional economic determinants of non-performing loans: Evidence from US states. *Journal of Financial Stability, 20*, 93–104.
Giesecke, K., Longstaff, F. A., Schaefer, S., & Strebulaev, I. (2011). Corporate bond default risk: A 150-year perspective. *Journal of Financial Economics, 102*(2), 233–250.
Goel, R. K., & Hasan, I. (2011). Economy-wide corruption and bad loans in banking: International evidence. *Applied Financial Economics, 21*(7), 455–461.
Gompers, P., & Lerner, J. (1998). Venture capital distributions: Short-run and long-run reactions. *The Journal of Finance, 53*(6), 2161–2183.
González-Coya, E., & Perron, P. (2024). Estimation in the presence of heteroskedasticity of unknown form: A Lasso-based approach. *Journal of Econometric Methods, 13*(1), 29–48.
Greenwald, B., & Stiglitz, J. E. (1987). Money, imperfect information and economic fluctuations. In *National Bureau of economic research (NBER) working paper, 0898–2937*. Retrieved from: https://www.nber.org/system/files/working_papers/w2188/w2188.pdf.
Hakura, D., & Cosimano, T. F. (2011). Bank behavior in response to Basel III: A cross-country analysis. In *International Monetary Fund (IMF) working papers* (pp. 1–34).
Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the Econometric Society, 47*(1), 153–161.
Hodula, M. (2022). Does FinTech credit substitute for traditional credit? Evidence from 78 countries. *Finance Research Letters, 46*, Article 102469.
Holland, B. (2021). *Is inflation returning? Not in sight but debate remains*. Bloomberg.com, 11 February. Retrieved from https://www.bloomberg.com/news/articles/2021-02-11/with-no-inflation-in-sight-why-the-inflation-debate-quicktake-kl0ods5m.
Jabeur, S. B., Gharib, C., Mefteh-Wali, S., & Arfi, W. B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change, 166*, Article 120658.
Jaffe, D., & Stiglitz, J. (1990). Credit rationing (Ch. 16). In *, Vol. II. Handbook of monetary economics*. Amsterdam: Elsevier Science Publication.
Jagtiani, J., & Lemieux, C. (2018). Do FinTech lenders penetrate areas that are underserved by traditional banks? *Journal of Economics and Business, 100*, 43–54.
Jakubík, P. (2006). Does credit risk vary with economic cycles? The case of Finland. In *, Vol. 11. Charles University in Prague–IES working paper*.
Jiang, C., Xiong, W., Xu, Q., & Liu, Y. (2021). Predicting default of listed companies in mainland China via U-MIDAS logit model with group Lasso penalty. *Finance Research Letters, 38*, Article 101487.
Johan, S., & Zhang, M. (2016). Private equity exits in emerging markets. *Emerging Markets Review, 29*, 133–153. https://doi.org/10.1016/j.ememar.2016.08.016
Jokivuolle, E., & Virén, M. (2013). Cyclical default and recovery in stress testing loan losses. *Journal of Financial Stability, 9*(1), 139–149.
Jones, S. (2017). Corporate bankruptcy prediction: A high dimensional analysis. *Review of Accounting Studies, 22*, 1366–1422.
Khieu, H. D., Mullineaux, D. J., & Yi, H.-C. (2012). The determinants of bank loan recovery rates. *Journal of Banking & Finance, 36*(4), 923–933.
Khrawish, H. A., Siam, W. Z., & Jaradat, M. (2010). The relationships between stock market capitalisation rate and interest rate: Evidence from Jordan. *Business and Economic Horizons, 2*, 60–66.
Kim, H., Cho, H., & Ryu, D. (2018). An empirical study on credit card loan delinquency. *Economic Systems, 42*(3), 437–449.
Kim, H., Cho, H., & Ryu, D. (2020). Corporate default predictions using machine learning: Literature review. *Sustainability, 12*. Art. 6325.
Kiyotaki, N., & Moore, J. (1997). Credit cycles. *Journal of Political Economy, 105*(2), 211–248.
Kukuk, M., & Rönnberg, M. (2013). Corporate credit default models: A mixed logit approach. *Review of Quantitative Finance and Accounting, 40*, 467–483.
Lawrence, E. C. (1995). Consumer default and the life cycle model. *Journal of Money, Credit and Banking, 27*(4), 939–954.
Li, E., Liao, L., Wang, Z., & Xiang, H. (2020). Venture capital certification and customer response: Evidence from P2P lending platforms. *Journal of Corporate Finance, 60*.
Liang, D., Lu, C. C., Tsai, C. F., & Shih, G. A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research, 252*(2), 561–572.
Lin, Y., Ye, Q., & Xia, H. (2023). Optimal interest rates personalization in FinTech lending. *Information Technology and Management*, 1–21.
Louzis, D. P., Vouldis, A. T., & Metaxas, V. L. (2012). Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios. *Journal of Banking & Finance, 36*(4), 1012–1027.
Ma, Q., Xu, L., Anwar, S., & Lu, Z. (2023). Banking competition and the use of shadow credit: Evidence from lending marketplaces. *Global Finance Journal, 58*, Article 100884.
McKinsey & Company. (2022). Europe's FinTech opportunity. October 26. Report. Available from https://www.mckinsey.com/industries/financial-services/our-insights/europes-fintech-opportunity.
Mei, Z., & Shi, Z. (2024). On LASSO for high dimensional predictive regression. *Journal of Econometrics, 242*(2), Article 105809.
Meinshausen, N., & Yu, B. (2009). Lasso-type recovery of sparse representations for high-dimensional data. *The Annals of Statistics, 37*(1), 246–270.
Menkhoff, L., Neuberger, D., & Rungruxsirivorn, O. (2012). Collateral and its substitutes in emerging markets' lending. *Journal of Banking & Finance, 36*(3), 817–834.
Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance, 29*(2), 449–470.
Merton, R. C., & Thakor, R. T. (2019). Customers and investors: A framework for understanding the evolution of financial institutions. *Journal of Financial Intermediation, 39*, 4–18.
Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing, 29*(1), 1–16.
Nam, C. W., Kim, T. S., Park, N. J., & Lee, H. K. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting, 27*(6), 493–506.
Nigmonov, A., & Shams, S. (2021). COVID-19 pandemic risk and probability of loan default: Evidence from marketplace lending market. *Financial Innovation, 7*(1), 83.
Nigmonov, A., Shams, S. M. M., & Alam, K. (2022). Macroeconomic determinants of loan delinquencies: Evidence from US peer-to-peer lending market. *Research in International Business and Finance, 59*, Article 101516. https://doi.org/10.1016/j.ribaf.2021.101516
Norden, L., & Wagner, W. (2008). Credit derivatives and loan pricing. *Journal of Banking & Finance, 32*(12), 2560–2569.
Norden, L., & Weber, M. (2009). The co-movement of credit default swap, bond and stock markets: An empirical analysis. *European Financial Management, 15*(3), 529–562.
Nourallah, M., Öhman, P., & Hamati, S. (2024). Financial technology and financial capability: Study of the European Union. *Global Finance Journal, 62*, Article 101008.
OECD. (2024). *"Population" (indicator)* (accessed on 12 October 2024) https://doi.org/10.1787/d434f82b-en.
Ogundimu, E. O. (2022). On Lasso and adaptive Lasso for non-random sample in credit scoring. *Statistical Modelling, 24*(2), 115–138.
Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 109–131.
Ooghe, H., & De Prijcker, S. (2008). Failure processes and causes of company bankruptcy: A typology. *Management Decision, 46*, 223–242. https://doi.org/10.1108/00251740810854131
Oren, R. (2013). The effects of usury laws: Evidence from the online loan market. *The Review of Economics and Statistics, 95*(4), 1238–1248.
Rivers, D., & Vuong, Q. H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics, 39*, 347–366.
Ronchetti, E., & Trojani, F. (2001). Robust inference with GMM estimators. *Journal of Econometrics, 101*(1), 37–69.
Roper, C. (2020). P2P lending looks to get credit moving during COVID-19 slowdown. *Yahoo! Finance*, 27 May, 2020. Retrieved from: https://finance.yahoo.com/news/p2p-lending-looks-credit-moving-033120299.html.

de Roure, C., Pelizzon, L., & Thakor, A. (2022). P2P lenders versus banks: Cream skimming or bottom fishing? *The Review of Corporate Finance Studies, 11*(2), 213–262.

Royston, P., & Lambert, P. C. (2011). *Flexible parametric survival analysis using Stata: Beyond the cox model*. College Station, Texas: Stata Press Publication.

Rubaszek, M., & Serwa, D. (2014). Determinants of credit to households: An approach using the life-cycle model. *Economic Systems, 38*(4), 572–587.

Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). Determinants of default in P2P lending. *PLoS One, 10*(10), Article e0139427.

Shin, K. S., Lee, T. S., & Kim, H. J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications, 28*(1), 127–135.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business, 74*(1), 101–124.

Sindreu, J. (2020). Why Covid-19 won't stop the new shadow banks. *The Wall Street Journal*, 13 July, 2020. Retrieved from: https://www.wsj.com/articles/why-covid-19-wont-stop-the-new-shadow-banks-11594639940.

Steijvers, T., & Voordeckers, W. (2009). Collateral and credit rationing: A review of recent empirical studies as a guide for future research. *Journal of Economic Surveys, 23*(5), 924–946.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review, 71*(3), 393–410.

Stiglitz, J. E., & Weiss, A. (1992). Asymmetric information in credit markets and its implications for macro-economics. *Oxford Economic Papers, 44*(4), 694–724.

Stulz, R. M. (2019). FinTech, bigtech, and the future of banks. *Journal of Applied Corporate Finance, 31*(4), 86–97.

Tanda, A., Schena, C. M., Tanda, A., & Schena, C. M. (2019). The regulatory framework and initiatives. In *FinTech, BigTech and Banks: Digitalisation and Its Impact on Banking Business Models* (pp. 83–100).

Tang, H. (2019). Peer-to-peer lenders versus banks: Substitutes or complements? *The Review of Financial Studies, 32*(5), 1900–1938.

Thakor, A. V. (2020). FinTech and banking: What do we know? *Journal of Financial Intermediation, 41*, Article 100833.

The Economist. (2021). Inflategate: How rising inflation could disrupt the world's economic policies. *The Economist*, 13 February, 2021. Retrieved from: https://www.economist.com/leaders/2021/02/10/how-rising-inflation-could-disrupt-the-worlds-economic-policies.

Tian, S., Yu, Y., & Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. *Journal of Banking & Finance, 52*, 89–100.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B: Methodological, 58*(1), 267–288.

Tibshirani, R. J., & Efron, B. (1993). An introduction to the bootstrap. *Monographs on Statistics and Applied Probability, 57*, 1–436.

Tomlinson, N., Foottit, I., & Doyle, M. (2016). A temporary phenomenon? Marketplace lending: An analysis of the UK market. London, UK. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/financial-services/deloitte-uk-fs-marketplace-lending.pdf.

Tyagi, S. (2022). Analyzing machine learning models for credit scoring with explainable AI and optimizing investment decisions. *arXiv*. https://doi.org/10.48550/arXiv.2209.09362. preprint arXiv:2209.09362.

Valla, N., Saes-Escorbiac, B., & Tiesset, M. (2006). Bank liquidity and financial stability. *Banque de France Financial Stability Review, 9*(1), 89–104.

Wadud, M., Ali Ahmed, H. J., & Tang, X. (2020). Factors affecting delinquency of household credit in the US: Does consumer sentiment play a role? *The North American Journal of Economics and Finance, 52*, Article 101132.

Wei, Z., & Lin, M. (2016). Market mechanisms in online peer-to-peer lending. *Management Science, 63*(12), 3999–4446.

Weiss, G. N., Pelger, K., & Horsch, A. (2010). *Mitigating adverse selection in P2P lending – Empirical evidence from prosper.com*. Available at SSRN 1650774.

Xia, Y., Li, Y., He, L., Xu, Y., & Meng, Y. (2021). Incorporating multilevel macroeconomic variables into credit scoring for online consumer lending. *Electronic Commerce Research and Applications, 49*, Article 101095.

Xie, X., & Zhu, X. (2022). FinTech and capital allocation efficiency: Another equity–efficiency dilemma? *Global Finance Journal, 53*, Article 100741. https://doi.org/10.1016/j.gfj.2022.100741

Yoon, Y., Li, Y., & Feng, Y. (2019). Factors affecting platform default risk in online peer-to-peer (P2P) lending business: An empirical study using Chinese online P2P platform data. *Electronic Commerce Research, 19*(1), 131–158.

Young, A. (2022). Consistency without inference: Instrumental variables in practical application. *European Economic Review, 147*, Article 104112.

Zhang, B., Baeck, P., Ziegler, T., Bone, J., & Garvey, K. (2016). Pushing boundaries: The 2015 UK alternative finance industry report. Cambridge, UK. Retrieved from https://www.jbs.cam.ac.uk/wp-content/uploads/2020/08/2015-uk-alternative-finance-industry-report.pdf.