June 2020

WORKING PAPER SERIES

2020-ACF-07

The Rise of Fintech: A Cross-Country Perspective

Oskar Kowalewski
IESEG School of Management, Paris & LEM-CNRS 9221, Lille, France

Paweł Pisany
Institute of Economics, Polish Academy of Sciences, Warsaw, Poland

IÉSEG School of Management Lille Catholic University 3, rue de la Digue F-59000 Lille Tel: 33(0)3 20 54 58 92 www.ieseg.fr

The Rise of Fintech: A Cross-Country Perspective

Oskar Kowalewski[a,b,c], Paweł Pisany[c,*]

[a]*IESEG School of Management, Paris, France*

[b]*LEM-CNRS 9221, Lille, France*

[c]*Institute of Economics, Polish Academy of Sciences, Warsaw, Poland*

Abstract

This study investigates the determinants of fintech company creation and activity using a cross-country sample that includes developed and developing countries. Using a random effect negative binomial model and explainable machine learning algorithms, we show the positive role of technology advancements in each economy, quality of research, and more importantly, the level of university-industry collaboration. Additionally, we find that demographic factors may play a role in fintech creation and activity. Some fintech companies may find the quality and stringency of regulation to be an obstacle. Our results also show the sophisticated interactions between the banking sector and fintech companies that we may describe as a mix of cooperation and competition.

Keywords: fintech, innovation, start up, developed countries, developing countries

JEL codes: G21, G23, L26, O30

[*] Corresponding author.

E-mail addresses: ppisany@inepan.pl (P. Pisany), o.kowalewski@ieseg.fr (O. Kowalewski)

1 Introduction

With the rapid development of new technologies, the habits and demands of the financial services customer is changing. According to Bain & Company (2017), customer preferences towards mobile, fast, and comfortable solutions is one of the most important drivers of financial technology (fintech) development. Thus, fintech is now a buzzword in the finance industry and in research. It is still an emerging industry and therefore even defining it still creates significant problems. The Financial Stability Board (FSB) presented the most popular definition, and the one we use in this study. It states that fintech is "technology-enabled innovation in financial services that could result in new business models, applications, processes or products with an associated material effect on the provision of financial services" (Financial Stability Board, 2017).

Fintech is reshaping the financial system globally as most new companies challenge and eventually usurp traditional financial services. In general, the creation of fintech companies results in faster and/or better financial services or services for an underserved segment. In other words, fintech companies should increase the efficiency of the financial system, which in turn should have a positive impact on long-term economic growth. Hence, the creation of fintech companies is crucial for financial system development. However, we still know little about fintech startups and what determines their establishment. This study, try to shed some light on the determinants that potentially boost fintech startup creation and help new innovative ventures both in developed and developing countries.

This study investigates the factors that may foster or hamper fintech startup formation. We are especially interested in the economic, financial, institutional, demographic, technology, and regulatory features of the environment that boost entrepreneurship and innovation in finance. We divide the sample and explore the differences between developed and emerging financial markets in terms of fintech formation determinants. As our goal is slightly different from than Haddad and Hornuf (2019), we do not pay attention to different segments of the fintech sector, but we treat it as one important economic phenomenon.

Using a cross-country sample of fintech startups, we explore the determinants of startups. In this study, we employ a random-effects negative binomial (RENB) model to analyze the creation and activity of fintech companies. Next, we employ machine learning (ML) algorithms with specific tools that allow us to explain and interpret the economic mechanism and interpretations, specifically interpretable ML, to study the determinants of fintech company creation further. We build this latter research framework with the following steps: (1) create models with a large number of explanatory variables, (2) assess the predictive power of the models, (3) indicate the features with the greatest impact on the predictions, and (4) indicate the direction of interlinkages between the number startups and the most important features.

Our results document the positive role of technology advancements in a particular economy, quality of research, and more importantly, the level of university-industry collaboration. Additionally, we show that demographic factors may play a role in fintech creation and activity. Some fintech companies may find that the quality and stringency of some regulations are an obstacle. Interestingly, we find a sophisticated interaction between the banking sector and fintech companies, which we may describe as a mix of cooperation and competition.

Our research contributes to the literature in two ways. First, our research contributes to the literature on the country-level factors that determine the creation of fintech companies. Our paper is closely related to Haddad and Hornuf (2019), who use a cross-country sample. We expand their model, however, by adding new variables and dividing the countries into developed and developing economies. Indeed, we document that some of the determinants are strongly related to the level of economic development.

Second, we extend the existing empirical work by employing ML algorithms that allow us to build models with a large number of explanatory variables and asses their importance in making predictions. In our opinion, the methods provide much better accuracy in terms of the determinants Moreover, we use additional tools to interpret our machine learning models and investigate the mechanisms behind fintech startup formation.

2 Literature review

Bömer and Maxin (2018) present an extensive literature review in which they distinguish four main topics that studies on fintech companies consider: 1) proper definitions framework, 2) fintech success factors, 3) legal aspects of fintech activity and regulatory challenges, and 4) relations between banks and fintech companies. Using Bömer and Maxin's (2018) classification of the fintech research, our study contributes to the second and fourth strands. Hence, we limit ourselves to discussing the literature on fintech related to those two strands.

The first strand of the literature focus on fintech success factors, which we can divide into two subgroups in our opinion. The first subgroup tries to determine the company-level factors that determine the success of a fintech company. According to Lee and Teo (2015), there are five key success factors, which they term LASIC as an acronym for the five factors: low margin, asset light, scalable, innovative, and compliance easy. The second subgroup relates to the environmental factors that may influence the success of a fintech company. Currently, there are few quantitative empirical studies on the determinants of fintech companies' creation at the country level. Moreover, most of these studies treat all countries as homogeneous. With our study, we aim to provide some new evidence on the creation and activity of fintech companies, both in developed and developing countries.

The FSB conceptual analysis (Financial Stability Board, 2019) indicates the supply side factors, such as technological advancements and regulatory schemes, as well as demand side factors; that is, changing customer expectations. In particular, among the vital technological advancements that create the opportunity for fintech development is mobile device availability. Haddad and Hornuf (2018) underline the importance of technology factors. They use a cross-country dataset to analyze the economic and technological determinants that induce entrepreneurs to establish fintech companies. They find that a country's development is positively related to economic development and the ready availability of venture capital. Additionally, they document that fintech development is positively correlated with the number of secure Internet servers, mobile telephone subscriptions, and an available labor force. Interestingly, they report that if it is more difficult for companies to access loans, then the country has a higher number of fintech startups. Overall, Haddad and Hornuf (2019) confirm that a set of economic and technology factors determine the fintech startup formation process.

The second strand of the literature analyzes the complex relationship between financial intermediaries, in particular, banks and fintech companies. Lacasse et al. (2016) underline the potentially disruptive power of fintech companies on financial intermediaries and discuss the fintech phenomenon as a competitive force against banks. Conversely, Holotiuk et al. (2018) show the potential motivation for and gains from bank-fintech company cooperation. In a complementary study, Jagtiani and Lemieux (2018) show that wide access to data and alternative data that may be used in credit scoring in the modern economy, can determine the creation of fintech companies. Buchak et al. (2017) show that a lack of regulation for new business models resulting in lack of capital and regulatory burdens can spur the growth of fintech companies.

Claessens et al. (2018) show that GDP per capita, the Lerner index in the banking sector, and the normalized regulation index, as well as dummies for each country (country effect) are significant drivers for fintech credit. They find positive signs on the relationships between total fintech credit and GDP per capita and the banking sector Lerner index. Moreover, the authors find a negative sign on the variable denoting the normalized regulation index and claim that more stringent financial regulation deters the development of fintech in the credit market.

The literature on fintech company development and its impact on the banking sector is still in development, and the results are often contradictory. With our study, we extend the understanding of the country-level factors in the creation of the fintech companies by distinguishing between developed and developing countries. Additionally, we hope to provide a better understanding of the interlinkages between fintech companies and other parts of the financial system, particularly financial intermediaries.

3 Data and methodology

3.1 Data

As we are interested in the country-level factors that determine the creation of fintech companies, our main dependent variable is the number of fintech companies established in a country in a given year. In the RENB estimation, we extend our analysis by employing an alternative dependent variable that represents the number of active fintech companies in a country in given year, and control for the number of closed fintech businesses.

We retrieve data on fintech companies from the Crunchbase platform, which contains detailed information on fintech startup formations, including investments and funding information. We collect data on the creation 4,708 fintech startups in 50 countries during 2005–2017. In the regression, we use the full sample as well as two subsamples consisting of developed (24) and developing countries (26), which we divide according to the MSCI stock market classification.[2] We believe that the stock market classification is a good proxy for the development of a country's financial system, which is crucial as we aim to study the interaction between fintech companies and financial intermediaries. In Attachment Table 1A, we provide the lists of developed and developing countries.

Table 1 shows the descriptive statistics for all variables used in the study. The data confirm significant variation in the scale of fintech startups and its development across countries and time. Hence, we believe that the database allows us to investigate the

[2]The classification criteria are available at https://www.msci.com/market-classification. The MSCI classification does not include Cyprus, Latvia, Luxembourg, which we added to the sample.

macroeconomic, financial, technology, social, demographic, and institutional features that foster or hamper fintech startup formation.

[Table 1]

In Attachment Table 2A, we present the description of all the variables used in the study. Overall, we employ 39 potential explanatory variables in the regressions, which represent a wide range of factors that determine the formation of fintech startups. We choose the variables based on the existing empirical research on fintech formation (Haddad and Hornuf, 2018), and retrieve them from the databases of the World Bank, International Communication Union, and Fraser Institute.

Figure 1 presents the descriptive statistics for all variables used, as well as the pairwise correlation matrix that is important to conduct a proper From-General-To-Specific procedure in the RENB estimation. To present a transparent correlation matrix, we use numeric labels for the variables. Table 1 provides the names of the variables corresponding to the numeric labels.

The explanatory variables include the economic development measure, GDP per capita (no. 4); one-period-lagged financing obtained by fintech companies (no. 3); and other determinants that can be divided into five broad groups: technology absorption in a country, financial system characteristics, education and science quality, economic freedom, and demographic structure. Figure 1 indicates an issue of multicollinearity between the variables in these groups, as well as between the features related to technology absorption and science development level across the countries in the sample. We control for them in our regressions by adding the variables one by one, as including significantly correlated regressors in one estimation may disturb the explanatory power of the model.

[Figure 1]

3.2 Methodology

In the study, we employ two distinct approaches to establish the country determinants of startup formation in the last decade. First, we use a traditional econometrics approach (Dushnitsky et al., 2016; Haddad and Hornuf, 2019), namely the RENB model. In the model, the dependent variable is the number of fintech companies established or active in country i in year t. In the regressions, the independent variables are lagged by one period to address the potential problem of reverse causality (Dushnitsky et al., 2016) and because interlinkages between variables may not necessarily be immediate. Lastly, we include year dummy variables to control for the year-specific effects.

Second, we employ the ML algorithms, which prior fintech studies do not use. Using ML algorithms allows us to build models with a large number of explanatory variables and assess their importance in making predictions. In this study with ML, we first employ a random forest (RF) model and then the extreme gradient boosting (XGB) model to compare prediction quality.

We can treat the RF as bagged decision trees. According to Breiman (2001), RF is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees. In other words, RF is a process to repeat a model with an additional random choice of predictors (regressors) considered at one node, which extends the effectives of regression trees significantly. Regarding the advantages of using RF, they can be used both in classification and regression tasks. Moreover, they are robust to the problem of correlated regressors. We can use a large number of potential explanatory variables without initial selection. Thus, researchers can also use RF as a tool to investigate the importance of explanatory variables in predicting the dependent variable and selecting regressors for the final model. Moreover, they reduce the problem of overfitting and are indifferent to non-linear interlinkages between the data (Kho, 2018; Walker, 2013; Drakos, 2019). Moreover, they are relatively easy to use due to the small number of parameters, such as *ntree* (number of trees to construct) and *mtry* (number of variables from the complete set of predictors considered creating nodes in a tree).

As an alternative ML algorithm, we employ XGB, which, like in the boosting procedure models, are created in an iterative way. These models are not built on random subsets of predictors (*mtry*) like in RF. Boosted trees are not created independently, but are determined by the errors of previous trees. In the boosting procedure, particular trees have an uneven impact on the final prediction. The idea behind this approach is that the algorithm has a chance to learn from its previous mistakes. Moreover, implementing regularization reduces the overfitting problem in XGB (Chen and He, 2019; Brownlee, 2016).

The most important drawback of RF and XGB relates to model interpretability (Kho, 2018; Walker, 2013). We overcome this problem by using the interpretable ML feature importance and partial dependence plots. Feature importance allows to indicate the list of the dependent variables with the most influence on the predictions of each ML algorithm. Partial dependence plots for the dependent variables allow us to establish the directions of interlinkages between fintech startups and the most influential factors (Molnar, 2019).

In general, we conduct our analysis using a sample for the period 2005–2017. However, introducing the lagged values of some variables reduces the sample period. This does not influence RENB model estimated on the unbalanced panel. However, we include the data for 2007–2017 in the ML models, as it creates an opportunity to account for the two lagged values of each variable (*t-1* and *t-2*). Moreover, we filled in the missing data in the ML algorithms using the median values for a particular country. This method should allow us to obtain a more accurate prediction of the number of newly created fintech companies in a given year.

4 Results

4.1 Main RENB results

Table 2 presents the main results using the RENB for the 50 countries in our sample. In columns (1)–(4), the dependent variable is the number of companies created, while columns (5)–(8) present the number of companies active in the market in country *i* in year *t*. We add the latter dependent variable to control for the number of active fintech companies in a country. In the sample, we correct the data for defaulted or acquired firms. The results in Table 2 show stable signs and sizes of the coefficients across all specifications.

Our results suggest that the structure and development of the country's financial system is an important determinant for fintech startup and activity. In columns (1) and (5), the coefficient of the access to loans variable is negative and significant for both dependent variables at the 1% level. The results suggest that fintech startups and activity are more likely to occur in economies with limited access to credit. Indeed, in columns (1)–(2) and (6)–(7), the coefficient of the variable representing non-performing bank loans is positive and significant, which is in line with our assumptions. Banks with increasing amounts of non-performing loans cannot provide households and enterprises with more debt financing, and are more likely to reduce lending. We assume, therefore, that financial innovation ventures form to fill the existing or growing gap in the financial system and as a tool for financial inclusion.

On the one hand, the coefficient of the Lerner index is positive and significant in all specifications with the number of active fintech companies as the dependent variable. These results may suggest that a banking sector with significant market power supports startup creation by collaborating with them, implementing their solutions, and introducing startup accelerators for fintech companies. On the other hand, the coefficients of bank credit are insignificantly related to the creation and activity of fintech companies. Thus, the level of development of the banking sector is not strongly related to fintech companies. Indeed, the coefficients of the bank-level net interest margin and return on asset (ROA) variables are also not significant in all specifications. Consequently, both the level of development and the profitability of banks seem not to determine the development of fintech companies.

In contrast to our expectation, the coefficient of the size of the equity market is insignificant in almost all specifications. As stock markets are associated with the financing of new technologies and risky companies, we assume a strong and positive relationship between the size of the equity market and fintech startups. Carpenter and Petersen (2002), using US data, show that most small high-tech firms obtain little debt financing. Hence, in their opinion, new equity financing, in the form of an initial public offering, is very important and permits a major increase in firm size. However, based on our results, this seems not to be the case for fintech startup financing and activity. One explanation could be that fintech companies often use private equity, especially crowdfunding, as a source of financing, and hence do not need a strong capital market for their development. Therefore, the European Commission is currently undertaking a new initiative to promote Europe as a more attractive location for start-ups by introducing new regulations on crowdfunding. The new regulation would enable crowdfunding platforms to access customers across the EU under one license issued by a single EU member state. It would also mean that the crowdfunding regulation would be similar to those of other financial intermediaries in the EU.

As expected, the variable representing alternative forms of financing, especially public grants available to fintech companies in the past, is positively related to the creation of fintech companies and the number of companies active in the market. The coefficient is positive and statistically significant at the 1% level. We assume that public grants are important for the creation of fintech companies, as the literature on entrepreneurship shows that access to financing is one of the most important challenges for startups (Egeln et al., 1997).

Overall the results indicate that fintech startups and activity is more likely to develop in a financial system with financial constraints, whereas fintech activity is more likely when strong banking groups are present. At the same time, the results show that fintech companies complement the existing financial intermediaries and its growing role has a positive impact on financial development, and hence for a country's economic growth in the long term. However, we find that public grants are important for the creation and development of new innovations, as fintech companies have very uncertain outcomes. The results are important from a policy perspective.

As expected, we find that the variables representing the quality and features of university and industry research play an important role for fintech startups and activity. The results show a strong positive relation between university-industry collaboration and startups and activity. The coefficient of the variable representing collaboration is significant at the 1% level in all specification. We find a positive and significant coefficient of the quality of research in the specification in which the dependent variable is active fintech companies. Consequently, the results indicate that effective collaboration between academia and business is more important for activity than the creation of fintech companies.

Interestingly, we find that the coefficient of legal enforcement is positive, while the coefficient of legal rights is negative. The coefficients of both variables are statistically significant in the specification using active fintech companies as the dependent variable. The result shows that institutions do not determine startups but only their operations. On the one hand, in line with the literature on law and financial development (Porta et al. 1997, 1998), we find that enforcement is important for the activity of the fintech companies. On the other hand, we find a negative relation between legal rights and fintech activity. One explanation is that the low level of legal right dampens the development of financial intermediaries, which in turn could be beneficial for fintech companies.

In contrast, we find that country-level development does not determine the creation or activity of fintech companies. However, we find that government size is positively related to the creation and activity of fintech companies, but the coefficient is only significant in the later specifications. In line with the literature, we find that the technology-related regressors are positively related to fintech development and activity (Haddad and Hornuf, 2019). The coefficients of the variables representing the usage of mobile-cellular telephones and fixed broadband subscriptions are positive and statistically significant at the 1% level.

Lastly, we find only weak evidence that demographic factors may determine the creation and activity of fintech companies. As expected, the coefficient of the variable representing the number of young persons in the population is positive and statistically significant, but only in two of four estimations. Younger people are often associated with high-risk profiles and are more likely to use fintech products (de Roure et al., 2016). In contrast, the coefficient of the share of older persons in the population is negatively related to the creation of fintech companies but positively related to their activity. The coefficient of the share of the urban population is not significantly related to the startup and activity of fintech companies. Hence, the relationship between demographic structure and fintech activity is ambiguous.

[Table 2]

4.2 Results for developed and developing countries

We hope to gain a better picture of the determinants driving the emergence of fintech companies by splitting the sample in two subsamples for developed and developing countries. Hence, we repeat our estimation using only the number of fintech startups as the dependent variable for the two subsamples. Columns (1)–(4) in Table 3 show the results for developed countries, while columns (5)–(8) show those for developing countries.

The results for the two subsamples are mostly in line with the results in Table 2 and confirm the positive and significant impact of the availability of financing for fintech ventures, the negative influence of rising access to loans, the positive impact of university-industry collaboration practices, and positive influence of technology-related factors.

More importantly, we find some interesting differences between developed and developing countries. The coefficient of the Lerner index is positive and significant in two estimations for developed markets, but negative in the estimation for emerging and frontier markets. We argue that this indirectly shows that in developed financial markets, banks with high market power are more likely to cooperate with fintech companies. We assume that well-established financial intermediaries do not see small innovative ventures as direct threats, but rather as potential sources of interesting ideas or consultancy to boost their own effectiveness. Thus, in developed markets, banks are interested in collaborating with fintech companies and even fostering their development by establishing specialized accelerators or startups. A good example of such a corporation is the digital bank Aion, whose development was backed by the private equity firm Warburg Pincus. The bank has a banking license in Belgium, which allows it to enter and operate in the European Economic Area (EEA) member countries.

In developing economies, financial intermediaries are weaker in general, and their reach is substantially smaller compared to developed markets. Moreover, the financial markets are still developing, and customers are beginning to learn about and use financial services. Hence, the financial intermediaries in such financial markets compete with fintech companies more directly.

In developed countries, however, the coefficient of non-performing loans is positively and significantly related to fintech development. The results suggest that banks are engaging in a riskier credit strategy as a response to fintech development. There are at least two possible explanations for this strategy in developed markets. First, banks may be forced to enter new, riskier markets in response to declining revenues due to fintech development in their home markets. Alternatively, banks may decide to enter riskier markets following fintech companies, which often serve markets financial intermediaries previously ignored. The increase in non-performing loans may also be the result of development of fintech by banks, which enter new markets and are experiencing a learning curve.

The results also suggest that a young demographic structure supports fintech formation in emerging markets in general. The coefficient of the share of young people in the population is positive, but statistically significant in one specification, while the coefficient of the share of old people in the population is negative, but also significant in only one specification. Interestingly, the results are quite the opposite for developed countries. The coefficient of the share of old people is positive and significant for developed countries, yet also statistically significant in only one specification, but at the 1% level. Moreover, the coefficient of the share of young people in the population is negative for developed countries, but statistically insignificant. Consequently, the results indicate that different customer groups are encouraging the development of fintech companies in developed and emerging markets. One explanation for the results is that in developed countries, the population is on average better educated and more accustomed in using IT technologies than in emerging markets. In developed countries, moreover, retirees represent a much larger group and are on average much wealthier and consequently more interested in financial services and products.

Additionally, we find that urbanization has some impact on the pace of fintech formation, but only in developed markets. One explanation for the result is that in emerging economies, fintech companies play an important role in providing people with basic financial services, especially in very remote places (fostering financial inclusion). Thus, in emerging and frontier markets, there is no relationship between financial innovation ventures and urbanization processes.

[Table 3]

4.3 Interpretable ML results

As a robustness check, we use interpretable ML to analyze the creation and performance of fintech companies. We conduct an RF and XGB analysis using the full sample as well as the two subsamples for developed and emerging markets.

In Table 4, we present the basic measures of model quality for the RF and XGB. The results indicate that the XGB model outperforms the RF model in predictions on the test data. XGB also offers higher R^2 values, particularly for the test data. Furthermore, the results above show that with our set of explanatory variables, we can explain and predict the number of fintech formations more accurately for developed markets than for emerging markets.

[Table 4]

As the XGB model provides better predictions than the RF model, we present only the results for the XGB model for brevity. The results for the RF model are available upon request. Additionally, we present panels of partial dependence plots to show the directions of the interlinkages between the prediction of the number of established startups in a given year and country, as well as the explanatory variables considered in the model.[3] We provide the partial dependence plots in the Appendix.

Figure 2 presents the permutation-based feature importance for the XGB model for the full sample. The results show that the most important variables that shape the environment for fintech is the lagged variable representing the value of financing for fintech companies. Moreover, the XGB model shows that university-industry collaboration in R&D is an important determinant of fintech formation. The results indicate that these two factors are the most likely to be the most important force driving the development of fintech in countries were academia and real economies work together on R&D.

In line with our previous results, we find that the availability of financing in the country's economy—debt, capital, or grants program support for innovative ventures—are important determinants explaining fintech creation in a country. We assume that the importance of this feature indicates some autoregressive elements in the phenomenon of fintech startup formation, as the financing companies obtained in previous years is obviously dependent on the number of the startups at that time.

Moreover, the results confirm the positive relation between the number of fintech company formations and the size of banking credit, as well as the ROA and net margin in the banking sector. We may explain these results in two ways. First, the banking sector is largely a recipient of fintech services, which offer services related to big data, for example. Good conditions and performance among banks results in higher demand for services from the

[3] We present Partial Dependence Plots for 20 regressors with the highest feature importance. The results using 30 variables are available upon request.

financial innovation sector, which in turn fosters the fintech formation process. While banks and fintech companies may be collaborators, there is also a significant area of competition between these two financial intermediaries. High ROA and net interest margins in the banking sector in a country may motivate smaller innovative entities to take on increased competition: offer better prices or look for niches. The results of the model also show that the relation between concentration in the banking sector and number of fintech companies is negative. We also observe a negative relation to the index of legal property rights protection, which may suggest that stringent regulatory frameworks still act as an obstacle for the fintech sector.

[Figure 2]

The methodology framework based on partial dependence plots also allows us to study the detailed shape of the relation between the predictions of the model and regressors, even for non-linear relationships. Figure 1A in the Appendix shows the partial dependence plots, from which we can confirm that fintech formation may increase sharply from improvements to university-industry collaboration, but only after reaching a certain, high threshold of that indicator.

We now present the results of the abovementioned analytical approach separately for the subsamples of developed and developing markets. The results in Table 4 suggest again that the XGB model explains the determinants of fintech formation in both markets to a greater extent.

Figure 3 shows the permutation-based feature importance for the XGB model for developed countries. In line with our previous results, we find that financing for fintech and university-industry collaborations are the main drivers of fintech formation in developed markets. Moreover, the net margin in the banking sector and the quality of research index are also important features that create a favorable environment for fintech startups. Hence, we find that the directions of the interconnections between these variables and the number of fintech formations are similar to those presented for the total sample.

In addition, we find that concentration and the lagged values of the Lerner index for the banking sector are relatively important features explaining fintech formation in developing countries. The results indicate that the higher the concentration and Lerner index, the higher the prediction of fintech startup formation. These results may suggest that banking sectors with significant market power support startup creation by collaborating with them, implementing their solutions, introducing startup accelerators, and sometimes acquiring fintech companies. Innovations in such financial systems due the market power of traditional intermediaries are ultimately located inside banks. In the future, the strong position and market advantage of banks, compared to their clients, may open up a path for increased competition from smaller innovative ventures. However, the relationships between the fintech sector and banks are sophisticated and may be described as a form of coopetition (collaboration with simultaneous competition). We argue that for now, the first explanation is more accurate.

To sum up, the financing-related, banking sector, and research-related variables seem to shape the conditions for fintech companies in developed markets. Technology, regulatory, and demographic issues are less important.

[Figure 3]

Figure 4 shows the permutation-based feature importance for the XGB model for the emerging markets. We find that the determinants of fintech startup creation in developing markets are different to some extent. In line with our results above, the results here confirm that the availability of financing for fintech ventures in the past plays a significant role. However, we find that the international trade proxy variable is now relatively important and influences the prediction of the number of fintech companies created, though negatively. One explanation for the result is that open international trade is highly correlated with foreign bank concentration, which in turn means higher competition. In other words, in more open markets, we may assume that fintech startups face increased competition from banks, especially foreign-owned ones. This competition in turn negatively effects fintech formation.

However, we find that the creation of fintech startups is positively interconnected with the time needed to enforce debt in emerging markets. Our results suggest that the more difficult it is to enforce debt in an economy, the more popular alternative financial services become. The negative relation with the variables denoting concentration in the banking sector and the quality of regulation is as expected. The results are interesting as the existing literature shows that credit rights and debt enforcement are positively related to banking credit (Bae and Goyal, 2009), and thus, to the development of the banking sector. Our results confirm that fintech companies and banks are more likely to be competitors in developing markets than in developed markets.

[Figure 4]

We also obtain interesting results for the regressor denoting the size (role) of the government in fintech formation in developing economies. Figure 3A in the Appendix illustrating the partial dependence plots shows a U-shaped relationship between the Fraser Institute indicator and the prediction of fintech startup formations. The lowest and highest values of this feature support the growth of fintech formations, suggesting that innovative companies arise in economies with both large and narrow government roles; however, pure types of economy-authority relations foster innovation in financial markets the most.

5 Conclusions

In this study, we investigate the determinants of fintech startup formation and potential differences between emerging and developed markets. Using data for 50 countries, we confirm the vital positive role of the availability of financing for the fintech sector, confirming Haddad and Hornuf's (2019) findings. Moreover, we document the positive role of mobile phones and fixed broadband subscriptions, which is in line with the analysis presented by the FSB (2019). More importantly, however, we show that the quality of research, and particularly university-industry collaboration, can contribute to the formation of innovative financial companies. Both of our analytical approaches—RENB and ML—confirm the role of financing availability and university-industry collaboration, as well as the robustness of our results.

Our results also show that regulatory factors are important. Buchak et al. (2017) indicate that the lack of dedicated regulation for new innovative business models is significantly shaping the fintech sector's growth prospects. In particular, we argue that while fintech startups are created, they cannot develop their activities because of regulatory burdens. On the one hand, our results indicate that a friendly legal environment helps innovation in financial markets. On the other hand, the regulator needs to be aware that some entrepreneurs may treat the fintech label as a tool for regulatory arbitrage while offering financial services. Ahern (2018) raises similar concerns, and we argue that this situation demands a proper balance between supervision and regulation. Moreover, the ML model results suggest that improving academia-business relations helps, but after reaching a certain threshold. Thus, it is worthwhile for policymakers to be patient in working on that aspect.

The results suggest that some factors driving the development of fintech companies are common between developed and developing countries. Among these factors, we find a positive impact of technology-related factors and university-industry cooperation on fintech development, and a negative impact of access to loans in a country. The last factor shows that fintech development may improve significantly with access to finance, which is important from a policy perspective.

On the other hand, our analysis reveals some differences in fintech sector development paths between the two countries groups. An important difference between those two markets is the relation between fintech companies and banks. Our results show that the interaction between the banking sector and fintech companies may be described as coopetition (mix of cooperation and competition), but with predominant cooperation in developed economies and predominant competition in emerging economies.

These results shed new light on the problem of bank-fintech company interconnections, as studied by Holotiuk et al. (2018) and Lacasse et al. (2016), among others. The coefficients of the Lerner index in our estimations lead us to conclude that in developed economies, banks with high market power treat fintech companies as a source of interesting ideas to boost effectiveness or mobile services quality. Thus, banks in developed markets support fintech startup formations, for example, by establishing accelerators. We claim that in emerging economies, more direct competition occurs. In line with this argument, the ML models show that the higher the Lerner index is and the higher the bank net margin and bank ROA in developed markets are, the higher the prediction of fintech startup formations. Moreover, the ML models suggest (in the sample of all 50 markets) that good performance among banks fosters fintech formation processes.

Lastly, we show that, in general, a young demographic structure supports the fintech company formation process. However, we find some differences between developed and developing countries. In developed countries, we find that an older population is positively related to the development of fintech companies. In our opinion, the results confirm that elder people in wealthier societies are willing to adopt innovative financial solutions. Urbanization has some positive impact on the pace of fintech formation, but only in developed markets. However, the ML models indicate that technology, regulatory, and demographic factors are relatively less important.

Based on our research, we make four policy recommendations. First, financing for innovative projects in the fintech sector is the key. Indeed, our results suggest that public policies are important, such as grants programs for small innovation companies. Second, public policy should pay attention to proper and effective communication between university and business. Some government-created coordination projects are advisable. Third, policymakers should not perceive fintech companies and banks purely as competitors, but rather foster accelerators organized by banks. Fourth, entrepreneurs may sometimes treat the establishment of fintech companies as a tool for regulatory arbitrage. This situation demands proper, balanced, and careful monitoring and supervision.

We are aware of the limitations of our study. In particular, we do not analyze the countries' legal environments in detail. The sample countries have different legal schemes and regulatory approaches to financial innovations, for example, the presence and practices of regulatory sandboxes. At the European Union level, the Payment Services Directive influences the fintech sector significantly as it enables certain non-financial entities (Third Party Providers) to have access, after permission, to the bank account data of selected clients. However, this would likely be a more suitable research area for in-depth legal studies. Moreover, we do not pay attention to technological details, for example, types of fixed broadband available in a country (based on cooper, coaxial, or fiber cables), which may significantly influence the effectiveness of the mobile infrastructure and, thus, the prospects for fintech startups. Furthermore, we concentrate mainly on fintech startups, but BigTechs (large technology companies such as Google, Amazon, Facebook, and Apple from the US, or Alibaba Group from China) also play a substantial role and are reshaping the financial industry. The abovementioned topics go beyond the scope of this study, and we leave them for future studies on financial innovations.

Acknowledgments

The authors would like to thank Adam Głogowski, Tomasz Gromek, Robert Jagiełło, and Tomasz Koźliński for their valuable comments and suggestions. Pawel Pisany acknowledges the help with machine learning methods at the Faculty of Economics, University of Warsaw, particularly from Piotr Wójcik, PhD.

Funding

This work was supported by the National Science Center (NCN), Poland, under grant no. 2019/33/B/HS4/00369.

References

Ahern, D. (2018). Regulatory Arbitrage in a FinTech World: Devising an Optimal EU Regulatory Response to Crowdlending. *EBI Working Paper Series no. 24*.

Bain & Company. (2017). Evolving the customer experience in banking. *Bain & Company and Research Now*.

Bae, K. H., Goyal, V. K. (2009). Creditor rights, enforcement, and bank loans. The Journal of Finance, 64(2), 823-860.

Bömer, M., Maxin, H. (2018). Why fintechs cooperate with banks—evidence from Germany. *Zeitschrift für die gesamte Versicherungswissenschaft, , Volume 107, Issue 4*, 359-386.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5-32.

Brownlee, J. (2016). *A Gentle Introduction to XGBoost for Applied Machine Learning*. Access at : https://machinelearningmastery.com/gentle-introduction-xgboost-applied-machine-learning/

Buchak, G., Matvos, G., Piskorski, T., Seru, A. (2017). Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks . *NBER Working Papers,* no 23288.

Campos, J., Ericsson, N. R., Hendry, D.F. (2005). General-to-specific Modeling: An Overview and Selected Bibliography. *Board of Governors of the Federal Reserve System, International Finance Discussion Papers,* 838.

Chen, T., He, T. (2019). *xgboost: eXtreme Gradient Boosting*. CRAN, mimeo.

Claessens, S., Frost, J., Turner, G., Zhu, F. (2018). Fintech credit markets around the world: size, drivers and policy issues. *BIS Quarterly Review*.

Carpenter, R.E., Petersen, B.C., 2002. Capital market imperfections, high-tech investment, and new equity financing. *Economic Journal*, 112(477), 54-72.

Drakos, G. (2019). *Random Forest Regression model explained in depth*, mimeo.

Dushnitsky, G., Guerini, M., Piva, E. Rossi-Lamastra, C. (2016). Crowdfunding in Europe: Determinants of Platform Creation. *California Management Review 58*, 44-71.

Egeln, J., Licht, G., Steil, F., 1997. Firm foundations and the role of financing constraints. *Small Business Economics*, 9(2), 137-150.

Financial Stability Board. (2017). *Financial Stability Implications from FinTech.*

Financial Stability Board. (2019). *FinTech and market structure in financial services: Market developments and potential financial stability implications.*

Fuster, A., Plosser, M., Schnabl, P., Vickery, J. (2018). The Role of Technology in Mortgage Lending. *Federal Reserve Bank of New York Staff Reports, no. 836*.

Greene, W. (2007). *Fixed and Random Effects Models for Count Data.* New York: Department of Economics, Stern School of Business, New York University.

Haddad, C., Hornuf, L. (2019). The emergence of the global fintech market: economic and technological determinants. *Small Business Economics, 53(1)*, 81-105.

Hausman, J.A., Hall, B.,Griliches, Z. (1981). Econometric models for count data with application to the patents-R&D relationship. *NBER Technical Paper Series*.

Holotiuk, F., Klus, M.F., Lohwasser, T.S., Moormann, J. (2018). Motives to Form Alliances for Digital Innovation: The Case of Banks and Fintechs. *BLED Proceedings, 22*.

Jagtiani, J., Lemieux, C. (2018). The Roles of Alternative Data and Machine Learning in Fintech Lending: evidence from the Lending Club Consumer Platform. *Federal Reserve Bank of Philadelphia Working Paper,* no 18-15.

Kho, J. (2018, October 19). *Why Random Forest is My Favorite Machine Learning Model*, mimeo.

Lacasse, R. M., Lambert, B., Osmani, E., Couture, C., Roy, N., Sylvain, J., Nadeau, F. (2016). A Digital Tsunami: FinTech and Crowdfunding. *International Scientific Conference on Digital Intelligence.* Quebec City, Canada.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R.W., 1997. Legal determinants of external finance. *Journal of Finance*, *52*(3), pp.1131-1150.

La Porta, R.L., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R.W., 1998. Law and finance. Journal of Political Economy, 106(6), 1113-1155.

Lee, D.K., Teo, G. S. (2015). Emergence of FinTech and the LASIC Principles. *Research Collection Lee Kong Chian School Of Business*.

Molnar, C. (2019). *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable.* Github, https://christophm.github.io/interpretable-ml-book/.

Philippon, T. (2015). Has the US finance industry become less efficient? On the theory and measurement of financial intermediation. *American Economic Review, 105(4),* 1408-1438.

Ribeiro, M. T., Singh, S., Guestrin, C. (2016). Why Should I Trust You? Explaining the Predictions of Any Classifier. *Proceedings of NAACL-HLT 2016*,. San Diego, California, 97-101.

Roure de, C., Pelizzon, L., Tasca, P. (2016). How does P2P lending fit into the consumer credit market? *Deutsche Bundesbank Discussion Papers, no 30/2016*.

Walker, M. (2013). *Data Science Central*, mimeo.

Zetzsche, D.A., Buckley, R.P., Arner, D.W., Barberis, J.N. (2018). From FinTech to TechFin: The Regulatory Challenges of Data-Driven Finance. *New York University Journal of Law and Business, 14(2)*, 393-446.

Table 1 Descriptive statistics

No.*	Variables	N	mean	sd	min	max
1	Fintech startups	650	7,243	24,69	0	276
2	Fintech active	650	47,39	182,7	0	2,260
3	Fin. fintech†	650	1,11E+08	7,05E+08	0	8,99E+09
4	GDPpc	650	29,429	24,023	868,9	111,968
5	Mobile	650	112,5	34,07	7,879	249,8
6	Internet	650	60,12	26	2,388	97,83
7	Fixed broad	647	19,53	12,61	0,00036	46,13
8	Branches	583	26,37	26,29	0,485	257,7
9	Accounts	148	78,26	23,97	9,72	100
10	Credit	586	96	52	10,65	253,3
11	Stock mkt	581	71,41	56,72	0,614	333,9
12	Net margin	598	2,875	1,99	0,125	10,49
13	ROA	596	0,9	1,099	-5,977	4,241
14	ROE	597	10,3	12,3	-101,5	38,47
15	Cost/Income	499	55,28	13,13	19,99	139,3
16	NPL	572	4,428	5,797	0,1	48,68
17	H-statistic	250	0,635	0,211	-0,0867	1,248
18	Lerner	463	0,265	0,132	-0,0842	0,939
19	Boone	495	-1,377	9,287	-101,2	2,814
20	Concentration	595	75,98	17,24	28,97	100
21	Legal rights	546	6,381	2,401	1	12
22	Scientists	550	4,626	0,604	3,089	6,297
23	Access to loans	550	3,578	0,932	1,58	5,744
24	Gov. procurement	550	3,837	0,615	2,466	5,552
25	Univ. collab.	550	4,387	0,813	2,427	5,968
26	Research	550	4,763	0,884	2,367	6,55
27	Absorption	550	5,286	0,624	3,634	6,461
28	Technology	550	5,523	0,83	3,331	6,875
29	Economic Freedom Index	600	7,305	0,745	4,493	9,199
30	Gov. size	600	6,187	1,256	3,249	9,41
31	Property rights	600	6,57	1,502	3,134	9,138

32	Sound money	600	8,78	1,155	3,253	9,922
33	International trade	600	7,626	0,914	3,6	9,603
34	Regulation	600	7,364	0,979	4,12	9,46
35	Enforcement	574	541,2	262,4	120	1,445
36	Age working	650	49,87	10,51	16,45	88,55
37	Pop. young	650	20,35	7,124	11,06	44,21
38	Pop. middle	650	67,04	4,618	53,04	85,87
39	Pop. old	650	12,61	5,61	0,75	27,05
40	Pop. urban	650	72,42	17,73	21,68	100
41	Pop. urban growth	650	1,433	1,642	-2,282	14,69

*These numeric labels correspond to those in the correlation matrix graph, Figure 1.

Figure 1 Correlation matrix

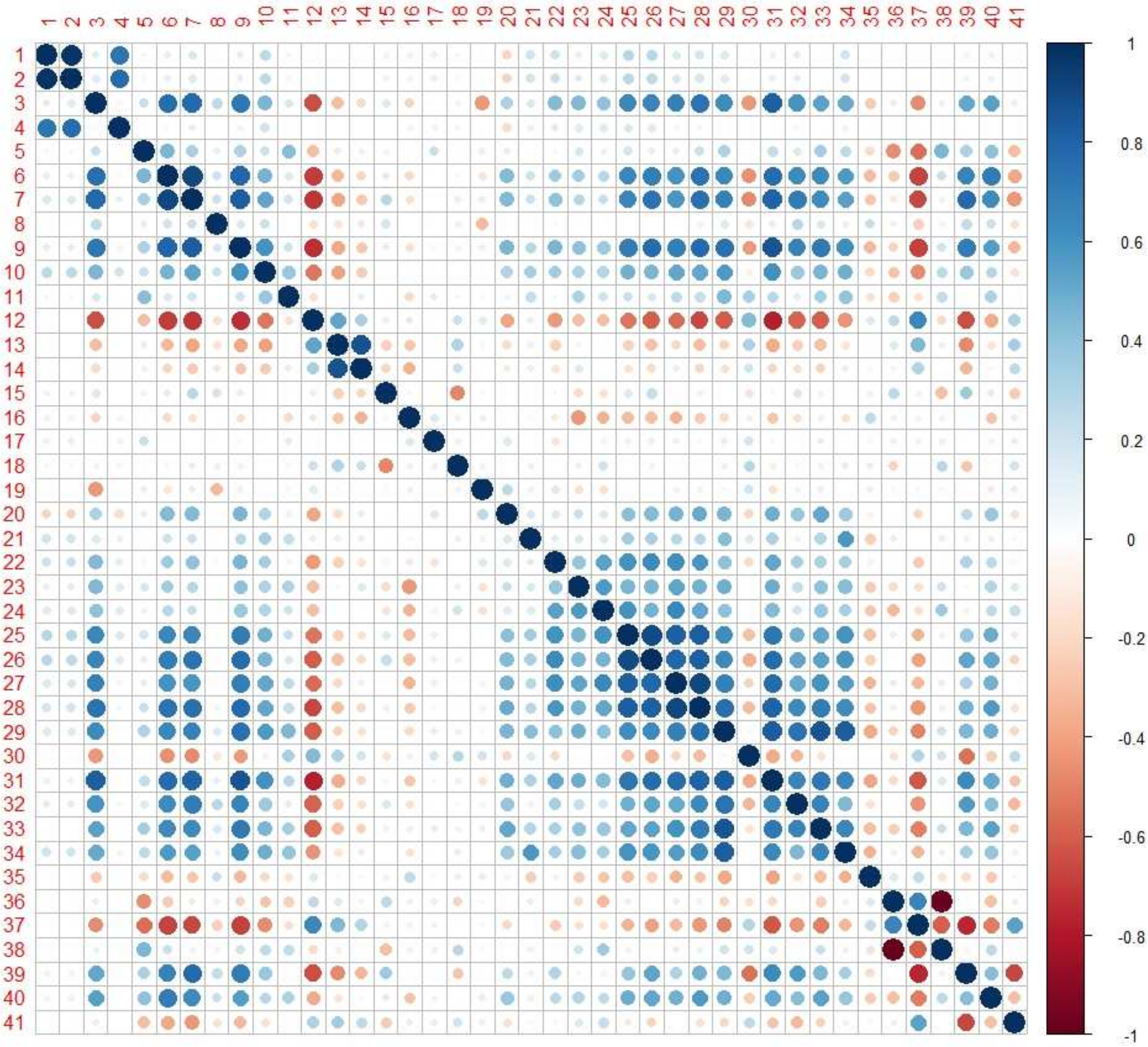


Table 2 Fintech startups and activities

This table reports the RENB regression results. In columns (1)–(4), the dependent variable is the number of founded fintech companies, and in (5)-(8), it is the number active fintech companies in country *i* in year *t*. All independent variables are one-period-lagged and their definitions are in Appendix in Table A2. The sample consists of 50 countries over 2007–2017. In all specifications, we use year dummy variables, but these are not reported for brevity. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at 1%, 5% and 10% level, respectively.

	Companies founded				Active companies			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Access Loans	-0.544***				-0.146***			
	(0.0949)				(0.0390)			
Credit		0.0048	0.0014	-0.0011		0.0003	-0.0015	0.0004
		(0.0031)	(0.0028)	(0.0019)		(0.0025)	(0.0023)	(0.0014)
NPL		0.0511**	0.0308*			0.0376***	0.0248***	
		(0.0211)	(0.0174)			(0.0093)	(0.0093)	
Lerner	1.960**	1.169	1.014		0.788***	1.474***	1.800***	
	(0.809)	(0.802)	(0.715)		(0.263)	(0.508)	(0.461)	
Net margin			0.0516				-0.0157	
			(0.0755)				(0.0503)	
ROA			0.108				0.0462	
			(0.0905)				(0.0476)	
Stock mkt	0.0025	-0.0011	-0.0028		-0.0000	-0.0012	-0.0042***	
	(0.0024)	(0.0025)	(0.0022)		(0.0011)	(0.0018)	(0.0016)	
Fin. fintech			0.0000***				0.0000***	
			(0.0000)				(0.0000)	

Univ. collab.			0.627***	0.476***			0.580***	0.345***
			(0.166)	(0.130)			(0.109)	(0.0945)
Research	-0.0523				0.248**			
	(0.206)				(0.0999)			
Scientists		-0.468**				-0.411***		
		(0.202)				(0.124)		
Enforcement	0.0006	0.0002			0.0008**	0.0001		
	(0.0006)	(0.0007)			(0.0004)	(0.0004)		
Legal rights				-0.0363				-0.0754***
				(0.0281)				(0.0211)
GDP			-0.0000				-0.0000	
			(0.0000)				(0.0000	
Gov. size	0.141				0.250***			
	(0.137)				(0.0658)			
Fixed broad	0.127***				0.115***			
	(0.0261)				(0.0130)			
Mobile		0.0141***		0.0210***		0.0173***		0.0191***
		(0.0052)		(0.00269)		(0.0027)		(0.0017)
Pop. young		0.0410		0.0419**		0.0241		0.0432***
		(0.0315)		(0.0192)		(0.0271)		(0.0139)
Pop. old	-0.0905**		0.0467		0.0900		0.0989***	

	(0.0428)		(0.0328)		(0.0624)		(0.0348)	
Pop. urban	-0.0201	0.0177		-0.0076	-0.00625	0.0072		-0.0101*
	(0.0132)	(0.0116)		(0.0075)	(0.0138)	(0.0102)		(0.0056)
Constant	1.979	-1.599	-3.075***	-3.494***	-0.0189	0.610	-1.643***	-2.161***
	(1.677)	(1.856)	(0.859)	(0.921)	(1.170)	(1.322)	(0.619)	(0.658)
Observations	289	280	328	482	289	280	328	482
Countries	49	49	49	50	49	49	49	50
AIC	1231	1234	1507	2353	1770	1840	2243	3704
BIC	1275	1278	1553	2391	1814	1883	2288	3742

Table 3 Fintech startups in developed and emerging markets

This table reports the RENB regression results. The dependent variable is the number of founded fintech companies in country *i* in year *t*. Columns (1)–(4) show the results for developed countries, while columns (5)–(8) show the results for emerging markets. All independent variables are one-period-lagged and their definitions are in Appendix in Table A2. The sample consists of 50 countries over 2007–2017. In all specifications, we use year dummy variables, but these are not reported for brevity. Robust standard errors are in parentheses. ***, **, and * denote statistical significance at 1%, 5% and 10% level, respectively.

	Developed markets				Emerging markets			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fixed broad	0.150***				0.118**			
	(0.0276)				(0.0534)			
Stock mkt	-0.0012	-0.0021	-0.0033		0.0052	0.0021	-0.0046	
	(0.0032)	(0.0031)	(0.0035)		(0.0047)	(0.006)	(0.0039)	
Lerner	2.069**	2.049	3.240***		-0.144	0.0351	-2.356*	
	(0.945)	(1.286)	(1.103)		(1.562)	(1.654)	(1.361)	
Access Loans	-0.436***				-0.527**			
	(0.113)				(0.259)			
Enforcement	0.0005	0.0030			-0.0002	0.0010		
	(0.0009)	(0.0026)			(0.0008)	(0.0009)		
Gov. size	0.321**				-0.233			
	(0.161)				(0.235)			
Research	0.0566				0.220			
	(0.288)				(0.419)			
Pop. Old	0.0324		0.229***		-0.208**		-0.0276	

	(0.0588)		(0.0876)		(0.0823)		(0.0384)	
Pop. Urban	-0.0022	0.0596*		0.00780	-0.0101	-0.00369		-0.0044
	(0.0253)	(0.0345)		(0.0155)	(0.0172)	(0.0207)		(0.0109)
Mobile		-0.00125		0.0165***		0.0227***		0.0248***
		(0.0116)		(0.0046)		(0.0063)		(0.0034)
Credit		0.00341	-0.0012	-0.0003		0.004	0.002	0.0026
		(0.00546)	(0.0052)	(0.0027)		(0.0074)	(0.0044)	(0.0039)
NPL		0.191***	0.128**			0.0224	-0.0044	
		(0.0427)	(0.0503)			(0.0226)	(0.0239)	
Scientists		-0.559*				-0.577		
		(0.290)				(0.355)		
Pop. Young		-0.147		-0.0147		0.0571		0.0479*
		(0.136)		(0.0484)		(0.0475)		(0.0262)
GDP			0.0000				0.0000	
			(0.0000)				(0.0000)	
Fin. Fintech			0.0000*				0.0000***	
			(0.0000)				(0.0000)	
Net margin			0.367*				0.00880	
			(0.188)				(0.0842)	
ROA			0.149				0.175	
			(0.159)				(0.129)	

Univ. collab.			0.539*	0.609***			0.730***	0.204
			(0.282)	(0.189)			(0.276)	(0.211)
Legal rights				-0.0589				0.006
				(0.0424)				(0.0381)
Constant	-3.986	-0.754	-6.894***	-4.030**	5.210*	-0.488	-1.488	-3.073**
	(2.775)	(3.327)	(1.823)	(1.598)	(2.873)	(3.255)	(1.142)	(1.398)
Observations	148	140	152	225	141	140	176	257
Countries	23	23	23	24	26	26	26	26
AIC	724.4	724.9	798.8	1304	504	501.2	691.3	1050
BIC	760.4	760.2	835.1	1335	539.4	536.5	732.5	1082

Table 4 ML models quality measures

	MSE	RMSE	MAE	MedAE	MSLE	R2
Panel A: Full sample (50 countries)						
RF model						
Train data metrics	123,2707	11,1027	4,1813	1,7121	0,5438	0,8702
RF model						
Test data metrics	42,5197	6,5207	4,2707	2,8058	0,8417	0,6430
XGB model						
Train data metrics	0,0015	0,0381	0,0053	0,0005	0,0000	0,9998
XGB model						
Test data metrics	25,8555	5,0848	2,8687	1,3884	0,4902	0,7829
Panel B: Developed countries						
RF model						
Train data metrics	152,256	12,339	5,9398	2,5932	0,7779	0,8929
RF model						
Test data metrics	299,048	17,293	7,15612	4,23067	1,21174	0,7537
XGB model						
Train data metrics	0,0051	0,0714	0,0141	0,0019	0,0000	0,9999
XGB model						
Test data metrics	112,539	10,608	4,897	2,321	0,656	0,9073
Panel C: Developing countries						
RF model						
Train data metrics	12,127	3,4824	1,7101	0,9782	0,3274	0,7232
RF model						
Test data metrics	24,8743	4,9874	2,4248	1,6077	0,5951	0,5610
XGB model						
Train data metrics	0,0148	0,1215	0,0223	0,0031	0,0001	0,9996
XGB model						
Test data metrics	17,7709	4,21555	2,26082	1,51517	0,53948	0,6864

Figure 2 XGB model

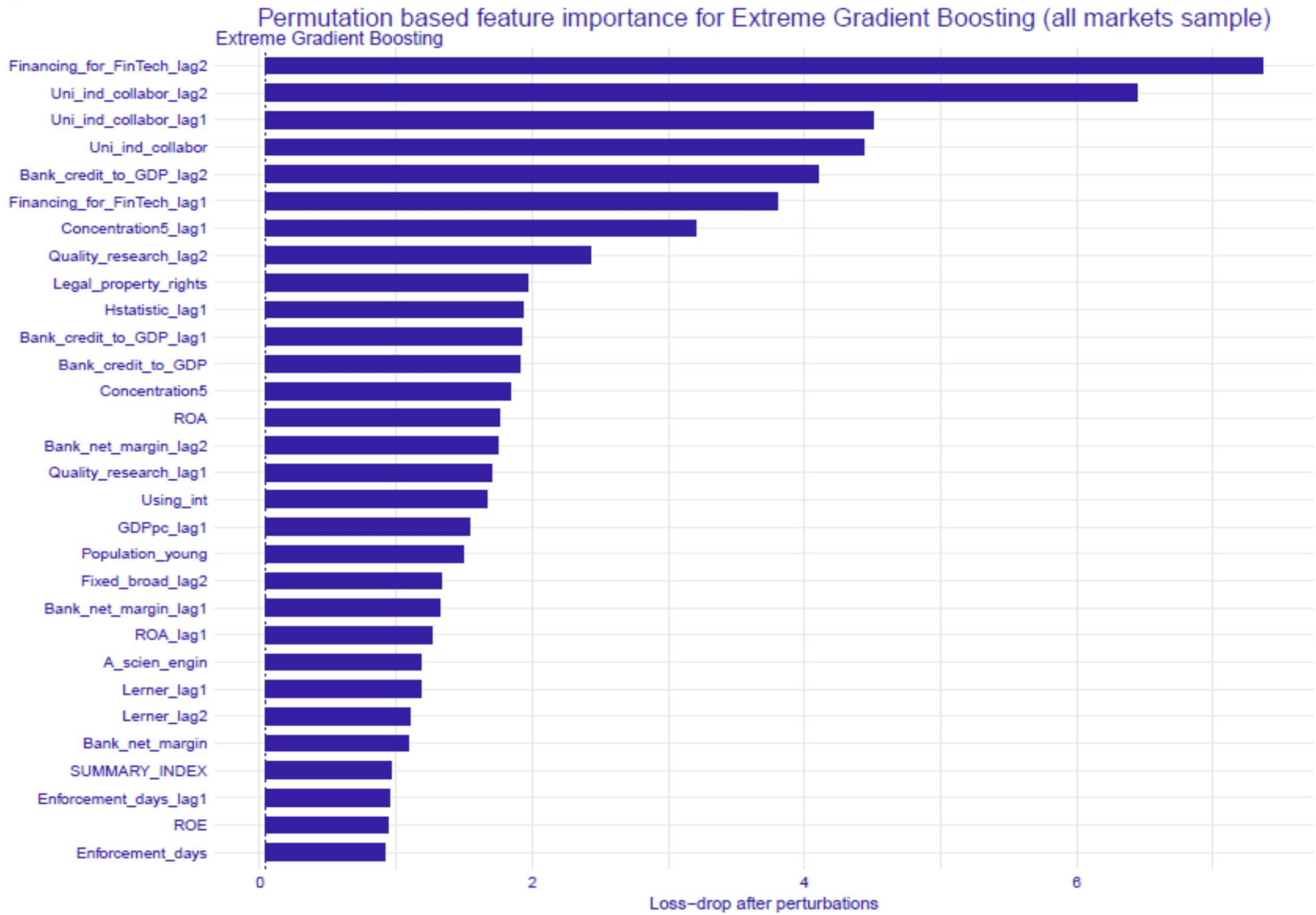


Figure 3 XGB model for developed countries

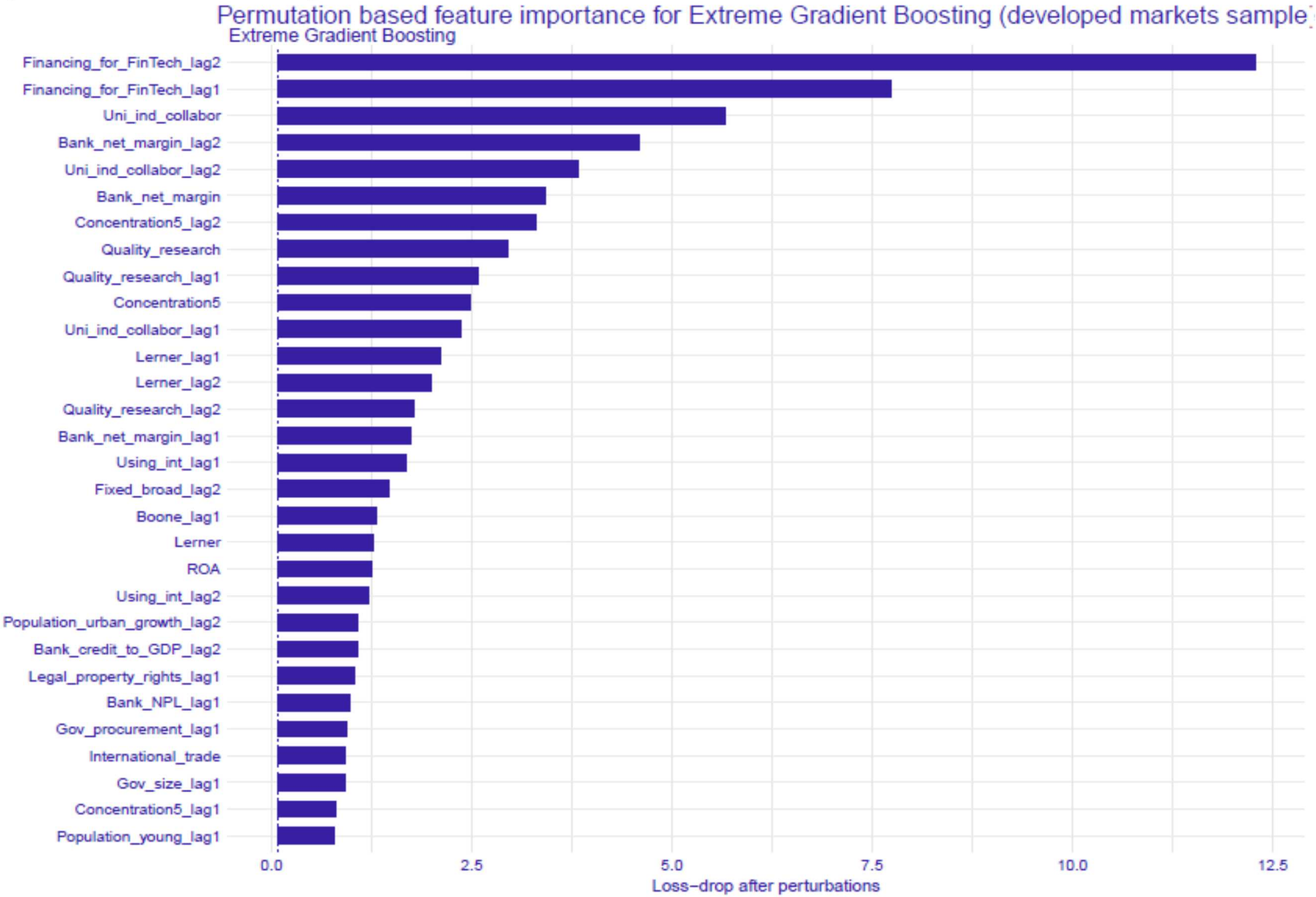
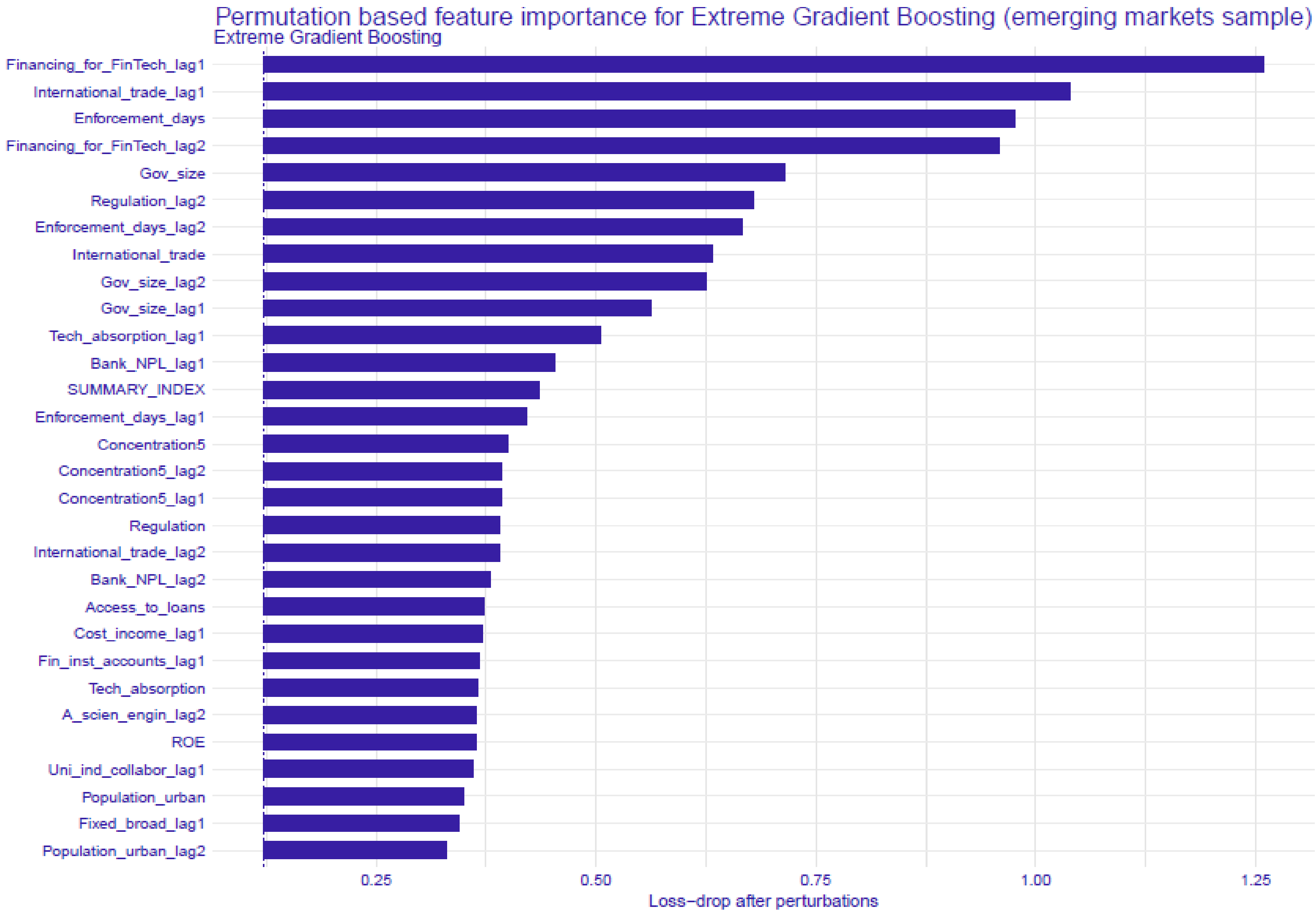


Figure 4 XGB model for developing countries



Appendix

Table A1 List of countries

Developed markets	Emerging markets
Australia	Argentina
Austria	Brazil
Belgium	Chile
Canada	China
Denmark	Colombia
Finland	Czech Republic
France	Cyprus
Germany	Estonia*
Hong Kong	Egypt
Ireland	India
Israel	Indonesia
Italy	Kenya*
Japan	Latvia
Luxembourg	Lithuania*
New Zealand	Malaysia
Norway	Mexico
Portugal	Nigeria*
Singapore	Poland
Spain	Russian Federation
Sweden	South Africa
Switzerland	South Korea
The Netherlands	Thailand
United Kingdom	Turkey
United States	Ukraine*
	United Arab Emirates
	Vietnam*

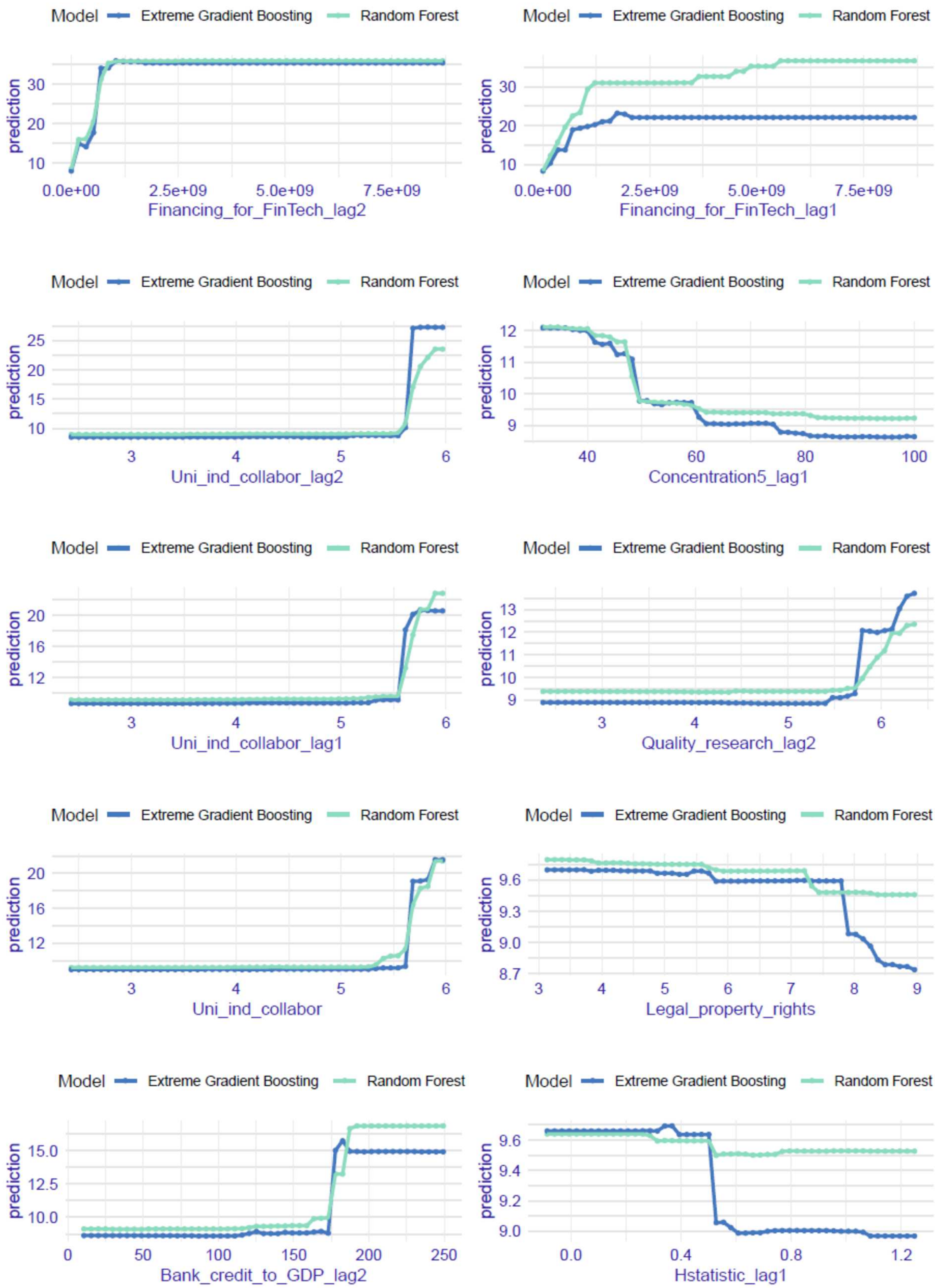
*Countries classified as frontier countries

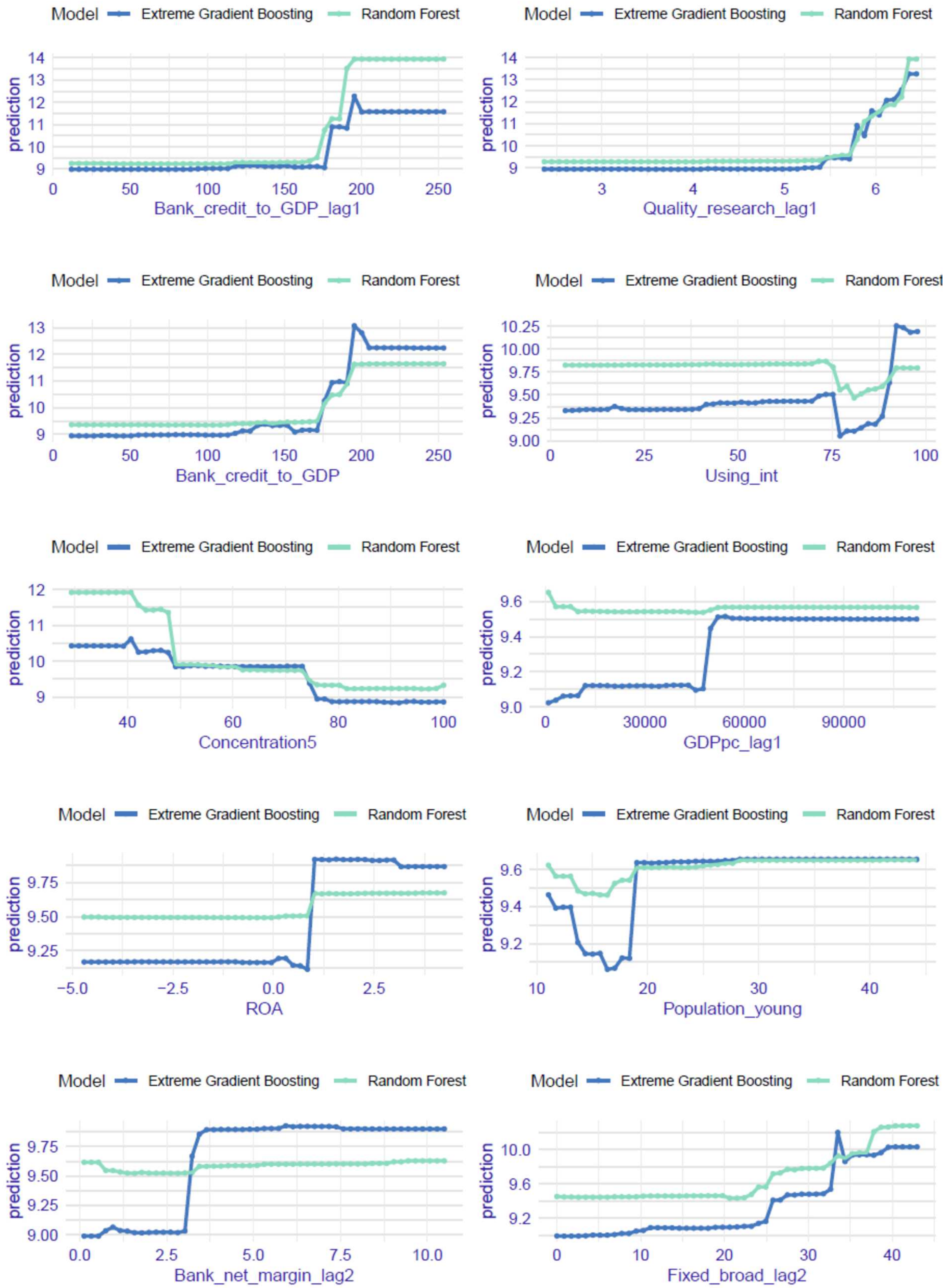
Table A2 Variable descriptions

Variable	Description	Source
Fintech startups	Number of companies founded in a given year	Crunchbase
Fintech active	Number of companies operating in a given year, still active companies, incl. information about closed fintech businesses	
Financing for fintech	Total sum of financing obtained by fintech companies in USD in a given year, equity and debt	
Mobile	Mobile-cellular telephone subscriptions per 100 inhabitants	International Communication Union
Internet	Percentage of individuals using the Internet	
Fixed broad	Fixed broadband subscriptions per 100 inhabitants	
GDP	GDP per capita (constant 2010 US$)	World Bank
Branches	Bank branches per 100,000 adults	
Accounts	Account at a formal financial institution (% age 15+)	
Credit	Domestic credit to private sector (% of GDP)	
Stock mkt	Stock market capitalization to GDP (%)	
Net margin	Bank net interest margin (%)	
ROA	Bank return on assets (%, after tax)	
ROE	Bank return on equity (%, after tax)	
C/I	Bank cost to income ratio (%)	
NPL	Bank nonperforming loans to gross loans (%)	
H-Statistic	H-statistic for the banking sector	
Lerner	Lerner index for the banking sector	
Boone	Boone indicator for the banking sector	
Concentration	5-bank asset concentration	
Legal rights	Legal rights index, 0–10 (best)	
Scientists	Availability of scientists and engineers, 1–7 (best)	Global Competitiveness Index
Access loans	Ease of access to loans, 1–7 (best)	
Gov. procurement	Government procurement of advanced technology products, 1–7 (best)	

Univ. collab.	University-industry collaboration in R&D, 1–7 (best)	
Research	Quality of scientific research institutions, 1–7 (best)	
Absorption	Firm-level technology absorption, 1–7 (best)	
Technology	Availability of latest technologies, 1–7 (best)	
Economic freedom	A summary index of the size of government, legal system and property rights, sound money, freedom to trade, and regulations	
Gov. size	Size of government	Fraser Institute Data
Property rights	Legal system and property rights	
Sound money	Sound money	
Trade	Freedom to trade internationally	
Regulation	Regulation	
Enforcement	Time required to enforce a contract (days)	
Age working	Age dependency ratio (% of working-age population)	
Pop. Young	Population ages 00–14 (% of total)	
Pop. middle	Population ages 15–64 (% of total)	World Bank
Pop. old	Population ages 65 and above (% of total)	
Pop. urban	Urban population (% of total)	
Pop. urban growth	Urban population growth (annual %)	

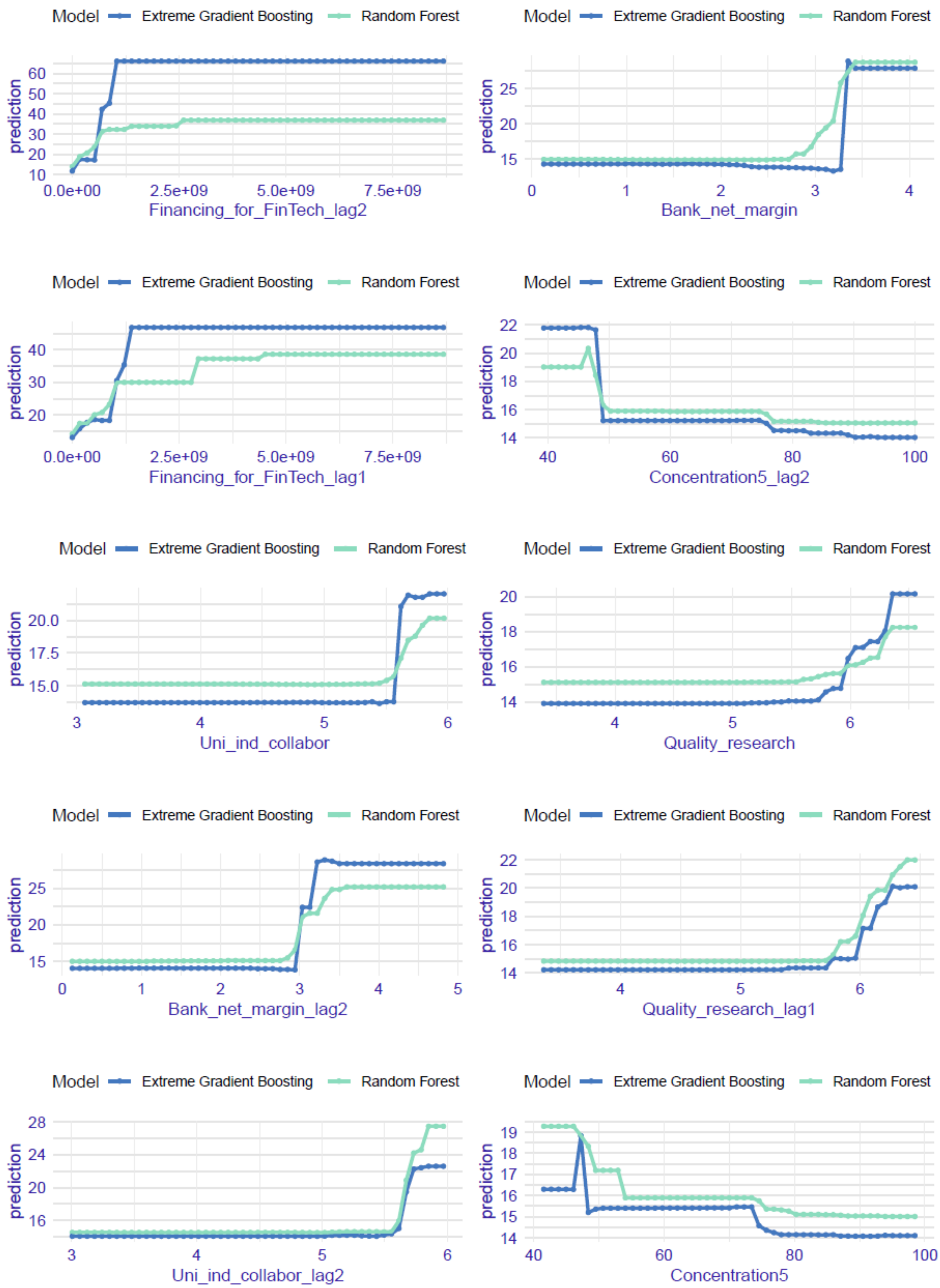
Figure 1A Partial dependence plots for the full sample (50 countries)





Model
Extreme Gradient Boosting
Random Forest
prediction
Bank_credit_to_GDP_lag1
Quality_research_lag1
Bank_credit_to_GDP
Using_int
Concentration5
GDPpc_lag1
ROA
Population_young
Bank_net_margin_lag2
Fixed_broad_lag2

Figure 2A Partial dependence plots for developed countries



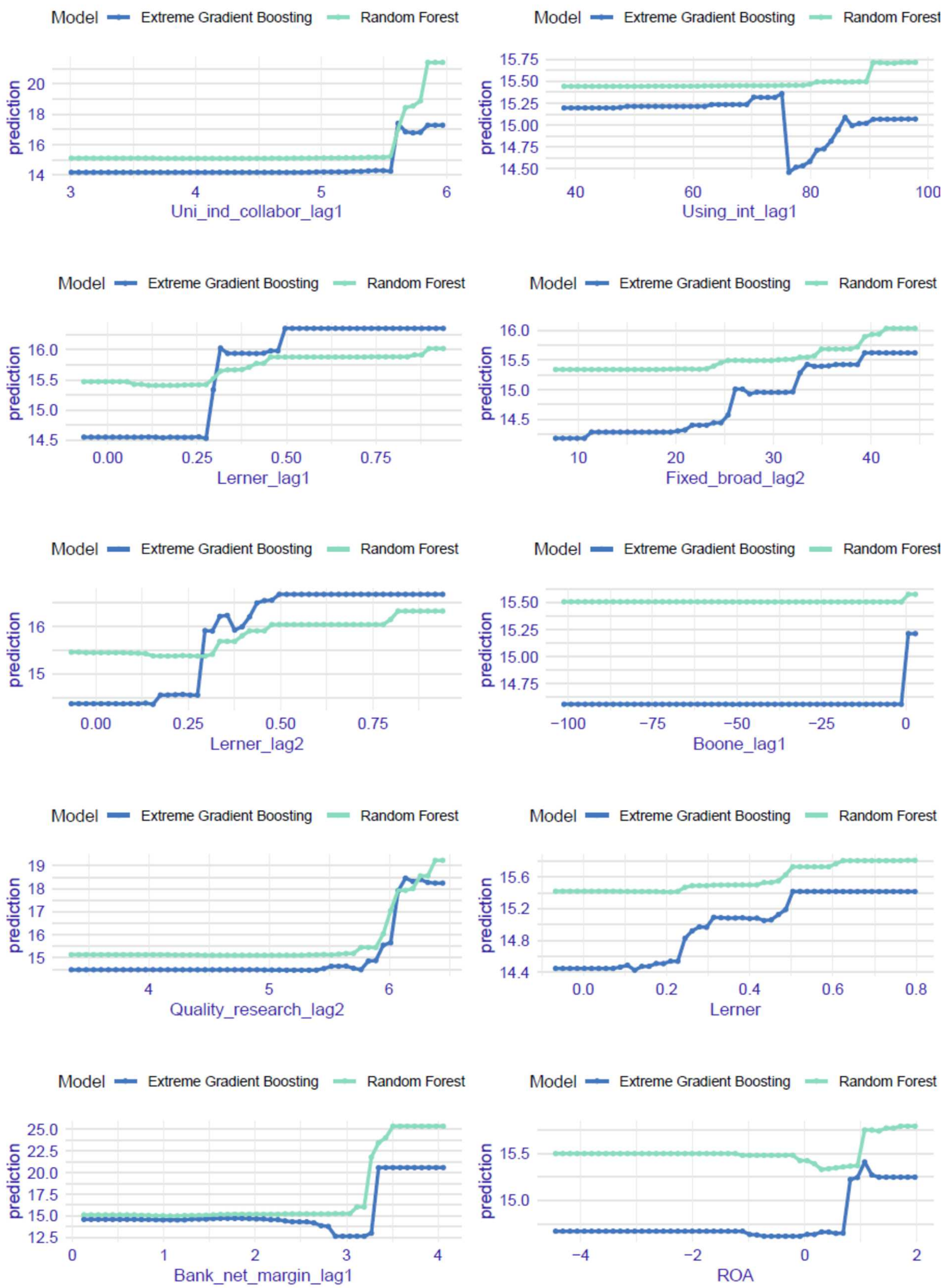
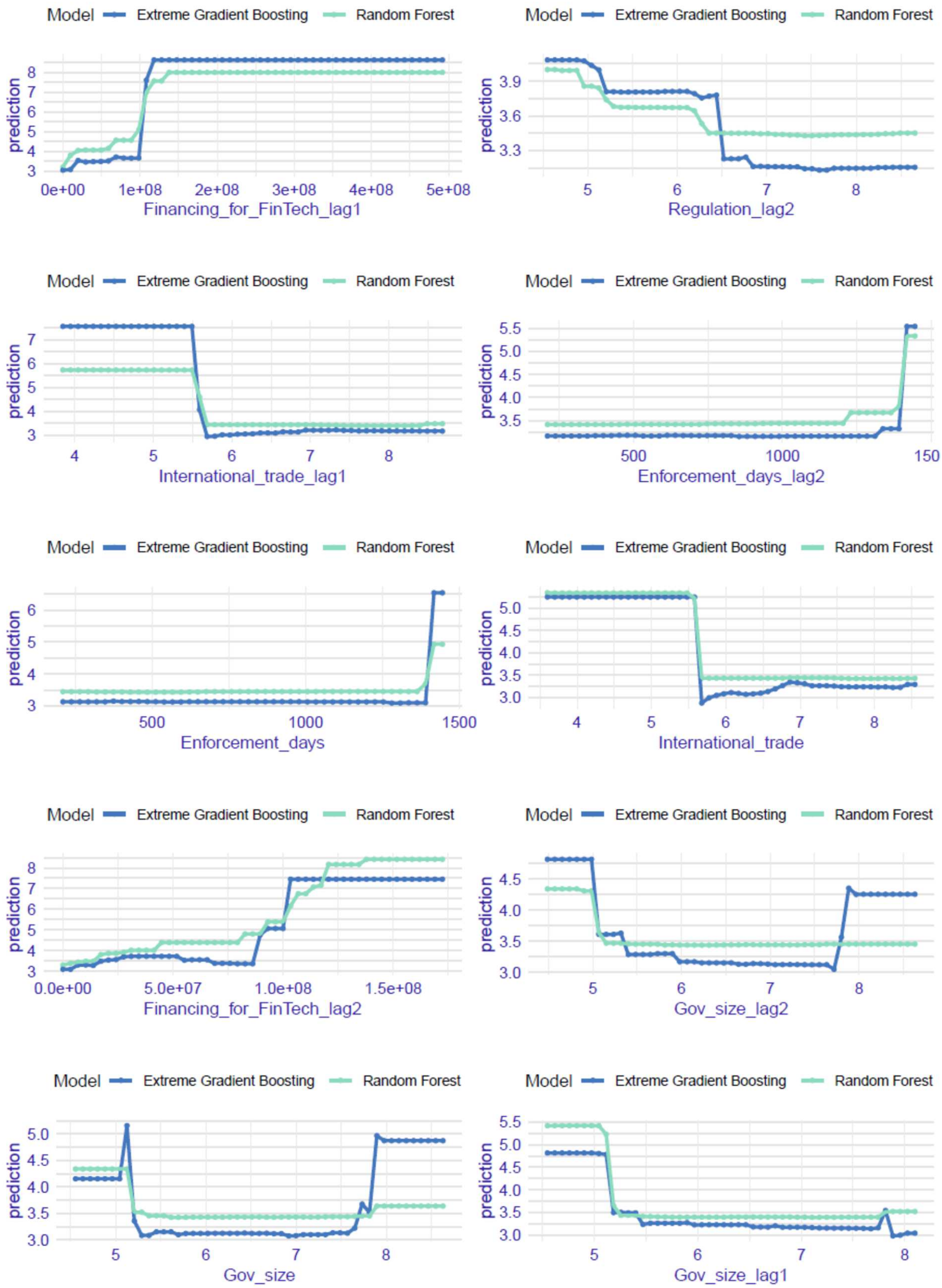
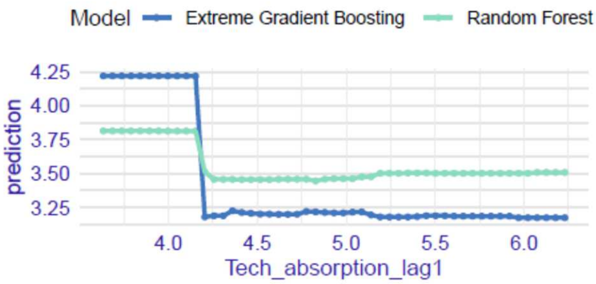
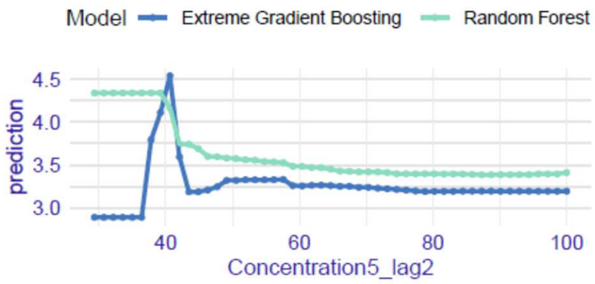
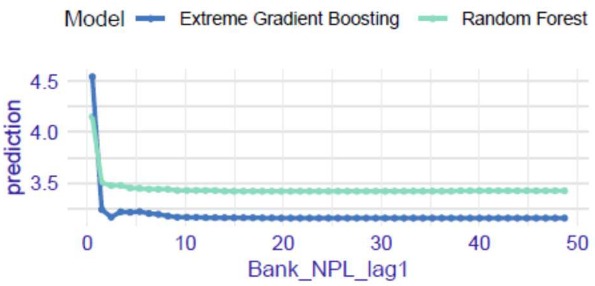

Model
Extreme Gradient Boosting
Random Forest
prediction
14
16
18
20
3
4
5
6
Uni_ind_collabor_lag1
14.50
14.75
15.00
15.25
15.50
15.75
40
60
80
100
Using_int_lag1
14.5
15.0
15.5
16.0
0.00
0.25
0.50
0.75
Lerner_lag1
10
20
30
40
Fixed_broad_lag2
15
16
Lerner_lag2
−100
−75
−50
−25
0
Boone_lag1
19
17
Quality_research_lag2
14.4
14.8
15.2
15.6
0.0
0.2
0.4
0.6
0.8
Lerner
12.5
15.0
17.5
20.0
22.5
25.0
1
2
Bank_net_margin_lag1
−4
−2
ROA

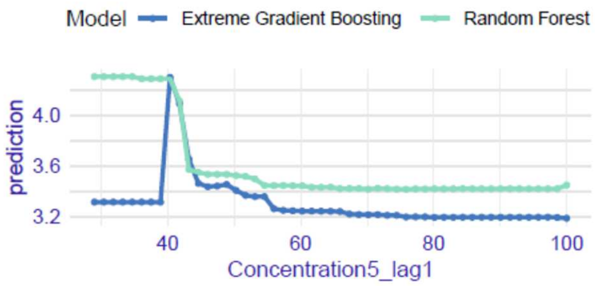
Figure 3A Partial dependence plots for developing countries

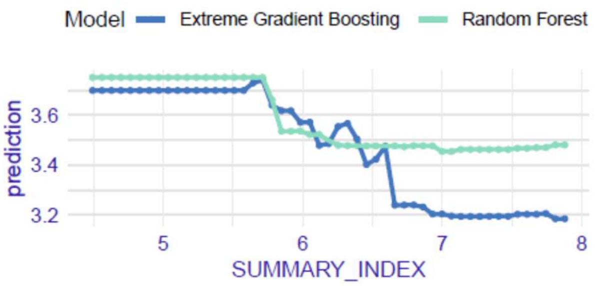


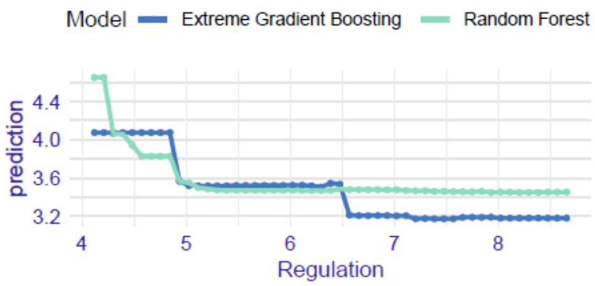

Model Extreme Gradient Boosting Random Forest
prediction
4.25
4.00
3.75
3.50
3.25
4.0
4.5
5.0
5.5
6.0
Tech_absorption_lag1

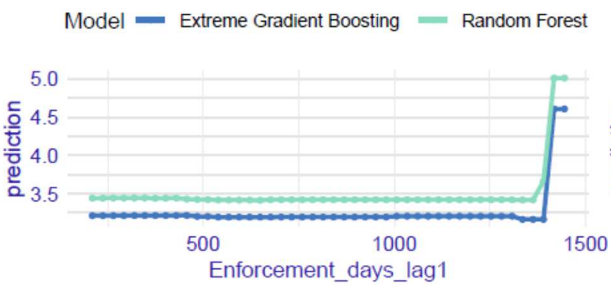

Model Extreme Gradient Boosting Random Forest
prediction
4.5
4.0
3.5
3.0
40
60
80
100
Concentration5_lag2

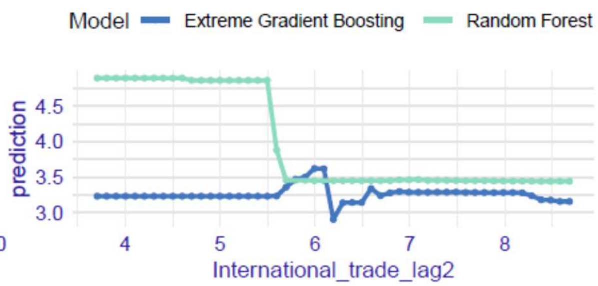

Model Extreme Gradient Boosting Random Forest
prediction
4.5
4.0
3.5
0
10
20
30
40
50
Bank_NPL_lag1

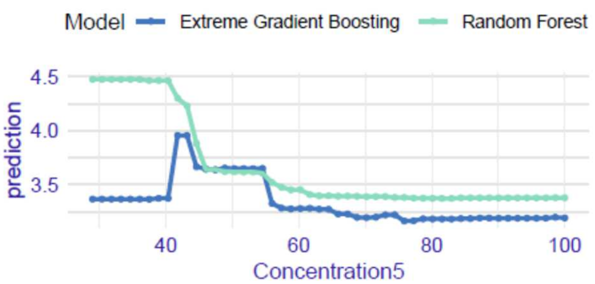

Model Extreme Gradient Boosting Random Forest
prediction
4.0
3.6
3.2
40
60
80
100
Concentration5_lag1

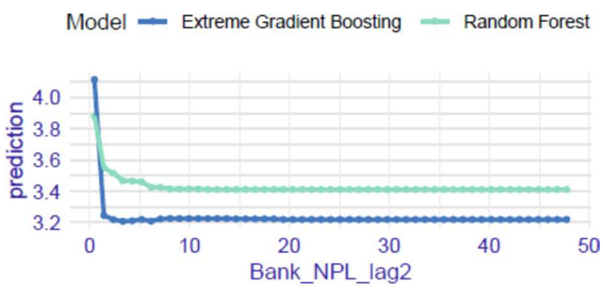

Model Extreme Gradient Boosting Random Forest
prediction
3.6
3.4
3.2
5
6
7
8
SUMMARY_INDEX


Model Extreme Gradient Boosting Random Forest
prediction
4.4
4.0
3.6
3.2
4
5
6
7
8
Regulation


Model Extreme Gradient Boosting Random Forest
prediction
5.0
4.5
4.0
3.5
500
1000
1500
Enforcement_days_lag1


Model Extreme Gradient Boosting Random Forest
prediction
4.5
4.0
3.5
3.0
4
5
6
7
8
International_trade_lag2


Model Extreme Gradient Boosting Random Forest
prediction
4.5
4.0
3.5
40
60
80
100
Concentration5


Model Extreme Gradient Boosting Random Forest
prediction
4.0
3.8
3.6
3.4
3.2
0
10
20
30
40
50
Bank_NPL_lag2