

Al-Maskati, Nawaf

Article

The determinants of bank profitability and risk: A random forest approach

Cogent Economics & Finance

Provided in Cooperation with:

Taylor & Francis Group

Suggested Citation: Al-Maskati, Nawaf (2022) : The determinants of bank profitability and risk: A random forest approach, Cogent Economics & Finance, ISSN 2332-2039, Taylor & Francis, Abingdon, Vol. 10, Iss. 1, pp. 1-15,
<https://doi.org/10.1080/23322039.2021.2021479>

This Version is available at:

<https://hdl.handle.net/10419/303558>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

The determinants of bank profitability and risk: A random forest approach

Nawaf Almaskati

To cite this article: Nawaf Almaskati (2022) The determinants of bank profitability and risk: A random forest approach, Cogent Economics & Finance, 10:1, 2021479, DOI: [10.1080/23322039.2021.2021479](https://doi.org/10.1080/23322039.2021.2021479)

To link to this article: <https://doi.org/10.1080/23322039.2021.2021479>



© 2022 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.



Published online: 05 Jan 2022.



Submit your article to this journal [↗](#)



Article views: 7699



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 11 View citing articles [↗](#)



Received 22 August 2021
Accepted 18 December 2021

*Corresponding author: Nawaf Almaskati, University of Waikato, Hamilton, New Zealand
E-mail: nawafmask@hotmail.com

Reviewing editor:
David McMillan, University of Stirling, Stirling, UK

Additional information is available at the end of the article

FINANCIAL ECONOMICS | RESEARCH ARTICLE

The determinants of bank profitability and risk: A random forest approach

Nawaf Almaskati^{1*}

Abstract: This study is the first to analyse the relative importance of a number of the most cited determinants of bank risk and profitability using random forest's relative value importance measure. The results show that a bank's profitability is largely determined by bank-specific factors, while a bank's risk is predominantly impacted by country-level factors. The results also suggest that proxies for market power and size play significant roles in impacting both the bank's profitability and its risk profile. The analysis also confirms the presence of a major role for a country's financial development status and regulatory quality in impacting the bank's riskiness. Lastly, the analysis confirms the presence of a small number of dominant determinants of a bank's profitability in contrast to the absence of clear dominant determinants of a bank's riskiness.

Subjects: Finance; Banking; Credit & Credit Institutions

Keywords: Bank performance; bank profitability; bank risk; bank stability; random forest; relative value

JEL Classification: G21; G28; E40

1. Introduction

The purpose of this study is to provide a comprehensive evaluation of the relative importance of the various determinants of a bank's profitability and risk. While a number of previous studies have attempted to study the various factors contributing to a bank's performance and/or risk (e.g., Bourke, 1989; Mirzaei et al., 2013; Molyneux & Thornton, 1992; Pasiouras & Kosmidou, 2007), we are aware of no study that attempts to study the relative importance of each factor. We use a well-known machine learning technique from other fields of research, namely random forest's (RF) relative variable importance (RVI), to measure the comparative significance of each factor in determining a bank's risk profile or profitability.

ABOUT THE AUTHOR

Dr. Nawaf Almaskati is an active researcher in financial markets, corporate governance and machine learning with several publications in the field. Dr. Almaskati's research focuses on empirically examining corporate governance and capital markets topics using various econometric and machine learning techniques to provide new insights into existing problems.

PUBLIC INTEREST STATEMENT

We find that a bank's profitability is largely determined by bank-specific factors, while a bank's risk is predominantly impacted by country-level factors. We also find that proxies for market power and size play significant roles in impacting both the bank's profitability and its risk profile. Furthermore, our analysis confirms the presence of a major role for a country's financial development status and regulatory quality in impacting the bank's riskiness. Lastly, our results suggest the presence of a small number of dominant determinants of a bank's profitability in contrast to the absence of clear dominant determinants of a bank's riskiness.

Our research adds to the growing stream of research focusing on understanding banks' performance and risk profiles through measuring and comparing the relative contribution of each individual variable to the process. It also provides equity and credit analysts with better insights on the importance of the role of each studied determinant of a bank's performance and riskiness. This will allow both groups of analysts to improve their analyses and forecasts of future performance through using findings on relative importance to assign unequal weights to the various factors used in their models based on importance.

The findings in this study also have some important policy implications. The insights from the new empirical evidence provided in this study provide regulators with a better understanding of the importance of each factor and help guide future regulatory decisions through focusing on those factors with the biggest impact. Similarly, the results from this study can help regulators monitor banks' performance more efficiently through directing their attention towards factors with higher contribution.

Using the random forest's relative value importance measure we find that the bank's profitability and risk are largely determined by bank-specific factors and country-level factors, respectively. We also find that market power and size proxies play an important role in impacting both the bank's profitability and risk. Lastly, the results also confirm a major role for a country's financial development status and regulatory quality in determining the bank's riskiness.

The remainder of this study is organized as follows. Section 2 provides a brief discussion of the past literature, while sections 3 and 4 explain the data collection and the methodology, respectively, including a short introduction to the random forest models and the relative value indicator. Section 5 discusses our results and findings, whereas section 6 provides some concluding remarks and suggestions for future research.

2. Literature review

A large number of studies have endeavoured to empirically examine determinants of a bank's performance and risk profile. Generally, a bank's performance is measured by the return on assets (ROA), return on equity (ROE) and net interest margin (NIM), whereas its risk profile is usually assessed by risk-weighted-assets (RWA) ratio, non-performing loans (NPL) ratio and the standard deviation of ROA or ROE (Li et al., 2021; Mirzaei et al., 2013; Molyneux & Thornton, 1992; Yanikkaya et al., 2018).

One of the first studies to examine determinants of a bank's profitability is Short (1979) who finds a significant role for market concentration, government ownership and assets growth in influencing a bank's profitability. A more comprehensive empirical assessment is provided by Bourke (1989) who examines the role of a number of internal (overhead expenses, liquidity, ownership and growth) and external (concentration, interest rates and inflation) factors in impacting a bank's performance as measured by the return on assets and return on equity. Bourke reports that the internal factors as well as concentration have a positive association with the bank's performance. A similar study by Molyneux and Thornton (1992) on a sample of European banks reports comparable findings. Molyneux and Thornton also find that banks with higher market power tend to exhibit higher risk avoidance.

More recent studies report significant roles for various bank-specific factors such as interest rates spread, non-interest income, off-balance sheet exposure, product diversification, loan provisions, and capital in determining the bank's profitability and risk profile. For instance, Goddard et al. (2004) report the existence of a significant positive relationship between capital to asset ratio and a bank's profitability, while Pasiouras and Kosmidou (2007) document the presence of a negative effect of cost-to-income ratio on a bank's performance in contrast to the positive

impact of capital adequacy ratios on the performance (see also, Athanasoglou et al., 2008; Berger & Bouwman, 2013). The relationship between capital and profitability is found to be directly linked to the bank's market power and ability to extract economic rent and operate more efficiently. Additionally, Berger and Bouwman (2013) document a significant role for capital in improving small banks' probability of survival during hard times.

In a related study, Mirzaei et al. (2013) observe that a greater market share in developed markets tend to lead to higher profitability, while also suggesting that policies targeted towards promoting competition may lead to destabilising the individual banks (see also, Saunders & Schumacher, 2000). They also report that higher interest-margin revenues in developing markets lead to more profitable and stable banks as banks use these higher margins to cover potential credit losses. In a global study covering 23 developed countries, Berger et al. (2009) report the presence of a negative relationship between market power and overall risk exposure. They also find evidence that while loan portfolio risk increases with market power the additional risk tends to be offset by the higher capital ratios. Further, Valverde and Fernández (2007) report that product diversification boosts profitability and improves market power as it compensates loss in interest-income due to increased competition and shrinking interest rate spreads. Additionally, Angbazo (1997) reports that off-balance sheet (OBS) exposure explains cross-sectional differences in interest rate risk and liquidity risk, which can be attributed to the off-balance sheet hedging activities. Angbazo finds that OBS activities help create a more diversified revenue generating base which reduces overall risk profile.

A number of studies also report the presence of a significant impact on a bank's performance and riskiness due to several country factors such as gross domestic product (GDP) growth, inflation, regulations, domestic credit and interest rate (Bourke, 1989; Molyneux & Thornton, 1992). For instance, Pasiouras and Kosmidou (2007) report that European banks' profitability is significantly impacted by GDP growth and inflation and that this effect differs for domestic versus foreign banks operating in the region (see also, Bolt et al., 2012; Dietrich & Wanzenried, 2014). Saunders and Schumacher (2000) find that regulatory policies along with macro interest rate volatility play a significant role in determining a bank's interest margin. They also document the presence of a trade-off between ensuring bank's solvency through imposing higher capital requirements and lowering the cost of banking services to consumers (i.e. lower interest margins). Similarly, Bolt et al. (2012) document a significant role for long-term interest rates in determining bank profitability during times of high economic growth. Further, Mirzaei et al. (2013) report the presence of a negative relationship between domestic credit as percentage of GDP and bank profitability in emerging markets versus a positive relationship in advanced markets. Mirzaei et al. also document that higher domestic credit leads to significantly higher bank risk in developing markets which could be attributed to the higher default probability and the less recovery potential in these markets (see also, Ash & Huizinga, 1999).

Previous studies use different statistical techniques to model the relationship between the bank's profitability or risk and its determinants. Earlier studies such as Bourke (1989) and Molyneux and Thornton (1992) employ simple linear regression models with variables such as ROA or ROE as the dependent variable and the rest of the determinants as explanatory variables. Later studies, utilize the General Methods of Moments (GMM) models on the basis that such models are able to better address some of the well-known problems present in such settings compared to other techniques such as fixed-effect models (Dietrich & Wanzenried, 2014; Maudos, 2017; Mirzaei et al., 2013; Pasiouras & Kosmidou, 2007; Tregenna, 2009). For instance, while fixed-effect models account for cross-sectional differences, they fail to account for potential endogeneity with regards to the dependent variable which is addressed by GMM models through employing a dynamic panel data approach (Yanikkaya et al., 2018). Moreover, a number of recent studies which focus on bank efficiency utilize various parametric (e.g., stochastic frontier approach) and non-parametric (e.g., data envelopment

analysis) frontier analysis methods to determine benchmark or best practice frontiers against which the performance of various banks is then measured (Claeys & Vander Venet, 2008).

3. Data

We obtain the entity-level data needed for our analysis from the Worldscope Database. The Worldscope Database contains detailed profile and financial data on public companies around the world. We extract year-end financial data for the period 2000 to 2019 for all active publicly traded banks with a minimum market capitalization of USD 500 million or equivalent as of the end of 2019.¹ The country-level data is obtained from the International Monetary Fund (IMF) and the World Bank databases. The various variables and their sources are explained in Table 1. Our final sample contains 1,245 banks from 66 countries.²

4. Methodology

4.1. Random forest and relative value importance

We use the relative value importance (RVI) indicator from the random forests (RF) model introduced by Breiman (2001) to assess the contribution of the various studied factors to a bank's risk and profitability. RF is an ensemble learning method that combines several random algorithms, or decision trees, to arrive at the final output. The RVI is calculated as the average weighted squared improvement to the model as a result of selecting the particular variable at each split (Friedman & Meulman, 2003). As highlighted by Biau and Scornet (2016), Hastie et al. (2009) and others, RF gained its popularity from requiring little intervention from researchers as well as its applicability to a wide array of classification and prediction tasks. Furthermore, RF models were found to be significantly less affected by everyday data challenges such as the presence of outliers and missing values (Biau and Scornet, 2016; Hastie et al., 2009). RF is also robust against many statistical issues impacting the performance of parametric models (e.g., regression analysis) such as multicollinearity or heteroskedasticity. Additionally, RF, like many other machine learning methods, is able to process a large number of input variables even in small samples, while also being largely unaffected by insignificant input variables (Hastie et al., 2009). Past studies have found that RF outperforms similar machine learning techniques such as decision trees and has comparable performance to other methods such as neural networks and generalized boosting (Jones et al., 2017).

4.2. Model specification and variables

The purpose of this study is to evaluate the relative importance of the different determinants of bank risk and profitability that were identified in past studies. To achieve this, we use a series of RF regressions with a number of profitability and risk proxies as dependent (output) variables:

$$\text{Profitability or Risk}_{i,t} = \text{Bank-level variables}_{i,t} + \text{Country-level variables}_{i,t} + \text{Year dummies} + \text{Country dummies} \quad (1)$$

Where all variables are defined in Table 1 along with a brief description of each variable and an explanation of the relationship with bank profitability and/or risk. Since risk and return are related with each indicator having an impact on the other, we include the profitability dependent variables (i.e. ROA and NIM) as independent variables in the models using a risk proxy as the dependent variable and vice versa (i.e. including risk-dependent variables as independent variables in the models using a return proxy as a dependent variable). We calculate the RVI values by assigning the most important variable a value of 100 and then re-express the values of all other variables on the same relative scale. We exclude the year and country dummies from the calculation of the RVI values as they are irrelevant to our study and have insignificant contributions to the percentage of variation explained in all cases (less than 2%).

Table 1. Variable definitions

Variable	Code	Category	Definition and Source	Explanation
Dependent variables				
Return on asset	ROA	Profitability	Net profit before tax/Total assets Source: Worldscope	A measure of the bank's asset profitability. A higher value indicates better profitability.
Net interest margin	NIM	Profitability	(Interest revenue—Interest expense)/Total assets Source: Worldscope	The spread between the bank's interest earnings and interest expenses as a proportion of its total assets. A higher value indicates a better performing bank.
Non-performing loans	NPL	Risk	Total non-performing loans/Total loans Source: Worldscope	A measure of the riskiness of the bank's loan portfolio. A larger ratio indicates a more risky portfolio. (Angbazo, 1997).
Z-score	ZSC	Risk	(ROA + Capital/Total assets)/ROA standard deviation for the last 5 years Source: Worldscope	A proxy for the bank's probability of failure. Larger values indicate less risk and better stability (Mirzaei et al., 2013)
Share price volatility	SVL	Risk	Volatility in share price over the relevant fiscal year Source: Worldscope	A market-based proxy for bank risk. A higher value indicates a more volatile share price and hence a higher risk.
Independent variables				
Bank-level				
Bank size	SIZ	Market power	Log of Total assets Source: Worldscope	A proxy for the bank's ability to benefit from economies of scale and market power (Goddard et al., 2004). Also, larger banks tend to have better survivability than smaller ones (Berger & Bouwman, 2013).
Capital ratio	CAP	Solvency	Capital/Total assets Source: Worldscope	A proxy for the bank's capital strength and solvency (Pasiouras & Kosmidou, 2007).

(Continued)

Table1. (Continued)

Variable	Code	Category	Definition and Source	Explanation
Lerner index	LRN	Market power	(Total income—Operating expenditures)/Total income Source: Worldscope	A measure of market power based on the bank's ability to set prices above its marginal cost. A larger index value is an indication of a higher ability to set prices above cost and thus a higher market power (Berger et al., 2009; Maudos & Solis, 2009).
Interest rate spread	IRS	Profitability	(Interest income/Total loans)—(Interest expense/Total deposits) Source: Worldscope	Measures the difference between lending and borrowing rates. A proxy for the bank's ability to earn interest income (Mirzaei et al., 2013).
Loans ratio	LON	Risk/Profitability	Total loans/Total assets Source: Worldscope	A measure of the bank's traditional activities (i.e. lending) as a proportion of total assets. A larger proportion of traditional lending activities may mean higher risk but also higher revenues (Claeys & Vander Vennet, 2008)
Non-interest income	NIC	Diversification	Non-interest income/Total assets Source: Worldscope	A proxy for the bank's revenue/product diversification. A larger ratio shows that the bank derives more of its revenues from non-traditional sources (i.e. other than lending interest; Maudos & Solis, 2009).
Operating expenditures	OPX	Efficiency	Operating expenditures/Total assets Source: Worldscope	This is a measure of the bank's efficiency in generating revenues from its expenditures. A larger value indicates lower efficiency (Mirzaei et al., 2013).

(Continued)

Variable	Code	Category	Definition and Source	Explanation
Risk-weighted-assets	RWA	Risk	Total risk-weighted-assets as defined by Basel Committee/ Total assets Source: Worldscope	A proxy for the riskiness of the bank's activities. It is also a measure of the cost of doing business in the form of capital requirements. A higher value indicates higher capital requirements and thus a higher risk and cost of conducting business (Saunders & Schumacher, 2000).
Off-balance-sheet activities	OBS	Profitability	Off-balance-sheet exposure/Total assets Source: Worldscope	A proxy to capture the bank's non-traditional revenue generating activities (Mirzaei et al., 2013).
Loan growth	LNR	Growth	Average inflation-adjusted growth in loans over 3 years Source: Worldscope	A proxy of the bank's growth and ability to generate profits (Mirzaei et al., 2013)
Liquidity	LIQ	Liquidity	Cash and marketable securities/Total assets Source: Worldscope	A measure of a bank's liquidity risk. It can also proxy for the cost of doing business as it captures the proportion of assets invested in low yielding assets to remain liquid (Berger & Bouwman, 2013).
Country-level				
Inflation	INF	Business cycle	- Annual percentage change in the consumer price index (CPI)	Inflation and GDP growth are measures of movement in the business cycle (Mirzaei et al., 2013).
Gross Domestic Product (GDP) growth	GDP		- Real annual growth in GDP Source: The World Bank	
Domestic credit	CRD	Financial structure	Domestic banking assets as percentage of GDP Source: IMF	A measure of the size of domestic credit market. A higher domestic debt to GDP may indicate a higher risk of default for banks (Mirzaei et al., 2013).

(Continued)

Table1. (Continued)

Variable	Code	Category	Definition and Source	Explanation
Regulatory quality	REG	Legal environment	A measure of the perception effectiveness and strength of local regulations Source: The World Bank	The index is provided by the World Bank and measures the perception of local government policies and regulations with regards to protecting and promoting private sector development. A higher value indicates better policies and regulatory environment.
Policy interest rate	POL	Financial structure	The average level of policy rate during the fiscal year as determined by the local central bank Source: The World Bank	A benchmark for the level of prevailing interest rates in the local market.
Financial development index	FDI	Market development	A measure of the development of the local banking sector and financial markets Source: IMF	The index is constructed by the IMF to rank countries based on the depth, access and efficiency of their banking sector and capital markets. A higher index value indicates a more developed market.

5. Results and discussion

5.1. Descriptive statistics and correlation analysis

Table 2 reports the summary statistics for the full sample. We can see that the values of some of the variables vary significantly which is expected given the period covered by our sample which includes the early 2000s recession related to the dot-com bubble, the global financial crisis in 2008 and the European debt crisis in 2010 in addition to several other smaller events.

In order to provide a preliminary assessment of the relationship between the various indicators, we report the correlation coefficients between the various dependent variables and the independent variables in Table 3. First of all, the vast majority of the correlation coefficients appear to be highly significant at the 1% significance level. Profitability measures (ROA/NIM) seem to be highly positively correlated with capital size (CAP), Lerner's index (LRN), interest rate spread (IRS), non-interest income (NIC), GDP growth (GDP) and policy rate (POL). Further, Lerner's index (LRN), financial development (FDI) and regulatory quality (REG) appear to be significantly negatively correlated with the bank's risk profile as measured by two risk indicators (NPL and SVL). They are also positively correlated with the third risk indicator (ZSC) which tends to have higher values for less risky banks. Many of these

Table 2. Descriptive statistics

Variable	Mean	Median	Std. Dev.	Min	Max
ROA	1.90	1.87	2.59	−52.53	22.60
NIM	0.03	0.03	0.02	−0.05	0.45
NPL	0.03	0.01	0.04	−0.16	0.81
ZSC	12.21	8.36	14.09	−6.04	178.25
VOL	30.11	26.83	14.24	2.03	156.01
SIZ	11.36	11.10	3.29	5.05	21.12
CAP	0.10	0.10	0.04	−0.12	0.54
LRN	0.21	0.23	0.34	−8.74	6.50
IRS	0.07	0.05	0.64	−17.16	28.78
LON	0.64	0.66	0.14	0.00	1.11
NIC	0.01	0.01	0.01	−0.03	0.18
OPX	0.02	0.02	0.02	0.00	0.28
RWA	0.56	0.63	0.37	0.00	7.10
OBS	0.47	0.12	3.48	0.00	139.40
LNR	0.11	0.07	0.17	−0.38	2.82
LIQ	0.06	0.04	0.06	0.00	0.70
INF	2.17	1.64	2.85	−3.75	17.15
GDP	5.84	4.06	5.79	−13.52	16.69
CRD	124.92	143.48	60.87	10.65	244.19
REG	0.83	1.26	0.79	−0.93	2.26
POL	2.92	1.63	3.63	−0.50	8.00
FDI	0.68	0.80	0.24	0.10	0.98

observations seem to be in line with prior findings in the literature. For instance, several studies have documented the presence of significant relationships between the bank's riskiness on one side and market power, local financial markets and regulatory environment on the other (Berger et al., 2009; Mirzaei et al., 2013; Saunders & Schumacher, 2000). Other studies have also linked indicators such product diversification, interest rate spread and business cycles to bank's profitability (Valverde & Fernández, 2007; Pasiouras & Kosmidou, 2007). While most of the reported correlations are highly significant, many of the correlation coefficients are relatively low especially in the cases including any of the risk indicators (NPL, ZSC and SVL). In the context of our study, this suggests that a bank's risk and profitability are likely to be impacted by a large group of variables as opposed to a few variables only. We examine such relationships in more details in the next section.

5.2. RVIs of the determinants of bank risk and profitability

We report the results of the RVI values from the various RF models in Figures 1–5 in order from the most important (top) to the least important (bottom) variable. Before discussing the RVI values, it is worth highlighting that three of the five models explain more than half of the total variance. The values of the percentage of variance explained are (the dependent variable is specified in the parentheses): 70.41% (ROA), 65.54% (NIM), 42.14% (NPL), 25.25% (ZSC) and 57.16% (SVL).

Figures 1 and 2 show the RVI values for the models containing our profitability indicators (ROA and NIM, respectively) as dependent variables. First important observation is that none of the top five

Table 3. Spearman's rank correlation

	ROA	NIM	NPL	ZSC	SVL
ROA	1.00				
NIM	0.58***	1.00			
NPL	0.07***	0.22***	1.00		
ZSC	0.47***	0.10***	-0.19***	1.00	
SVL	-0.15***	0.16***	0.24***	-0.42***	1.00
SIZ	0.16***	0.34***	0.11***	-0.03	0.07***
CAP	0.50***	0.53***	0.12***	0.13***	-0.08***
LRN	0.74***	0.12***	-0.14***	0.54***	-0.32***
IRS	0.44***	0.65***	0.31***	0.07***	0.18***
LON	-0.02	0.15***	-0.27***	0.04**	-0.11***
NIC	0.41***	0.33***	0.35***	0.02	0.02
OPX	0.29***	0.72***	0.22***	-0.10***	0.17***
RWA	0.25***	0.39***	0.00	0.11***	-0.01
OBS	0.17***	0.09***	0.16***	0.05***	-0.05***
LNK	0.36***	0.16***	-0.04**	0.16***	-0.02
LIQ	0.05***	-0.01	0.27***	-0.02	0.09***
INF	0.25***	0.36***	0.11***	0.05***	0.11***
GDP	0.38***	0.42***	0.21***	0.04**	0.14***
CRD	-0.11***	-0.03*	0.49***	-0.13***	0.13***
REG	-0.16***	-0.18***	-0.45***	0.08***	-0.23***
POL	0.44***	0.43***	0.27***	0.06***	0.07***
FDI	-0.18***	-0.13***	-0.55***	0.11***	-0.17***

The table shows Spearman's rank correlation coefficients between each of the dependent variables (columns) and all other variables (rows). *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level.

determinants of ROA are country-level variables which indicates that ROA is largely determined by bank-specific characteristics and performance. Next, it is also important to note that Lerner's index (LRN), a proxy of market power, has almost twice as much importance in determining a bank's ROA as the next best indicator, ZSC, which is an indicator of a bank's risk profile. This observation is supported by past findings in the literature regarding the role of market power in determining the bank's profitability (Berger et al., 2009; Maudos & Solís, 2009). The rest of the indicators show that bank risk profile (ZSC), operating efficiency (OPX), product diversification (NIC) and capital strength (CAP), play an important role in determining a bank's ROA. Further, the five most important determinants of a bank NIM are (in order of importance): interest rate spread (IRS), operating efficiency (OPX), policy interest rate (POL), financial development index (FDI) and size (SIZ). Once again, albeit it is now the first two variables rather than the first one only, IRS and OPX, are twice as important as the next best variable, POL, in determining the bank's NIM. Also, as expected, two of the top five indicators are related to interest rate (IRS and POL). Interestingly, the results show OPX as one of the top five indicators again, which confirms that a bank's efficiency with regards to its expenditures plays a significant role in impacting its performance as measured by both ROA and NIM. Further, the results also indicate that a bank's NIM is largely dependent on the development status of the market in which it operates as well as market power as measured by FDI and SIZ, respectively. A highly developed financial market is more likely to offer cheaper and more diversified funding resources for banks which explains the important role played by such indicator. Moreover, a bank's size and market power will determine its

Figure 1. Relative Score
(Dependent variable: ROA).

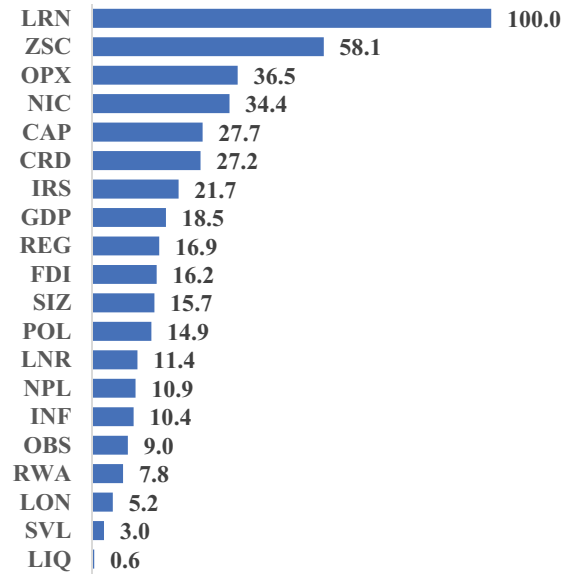
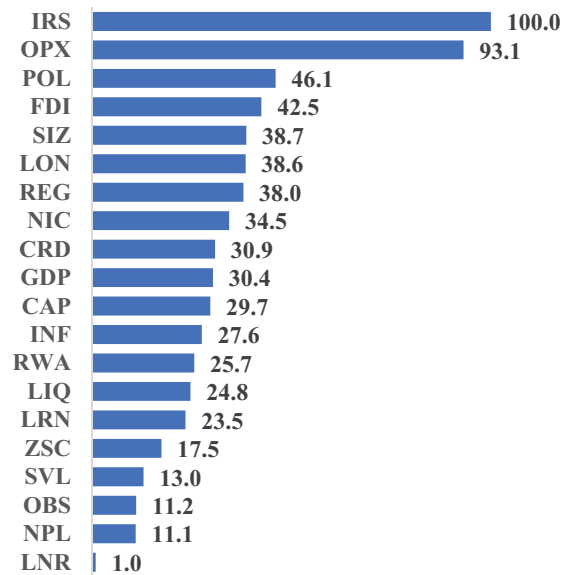
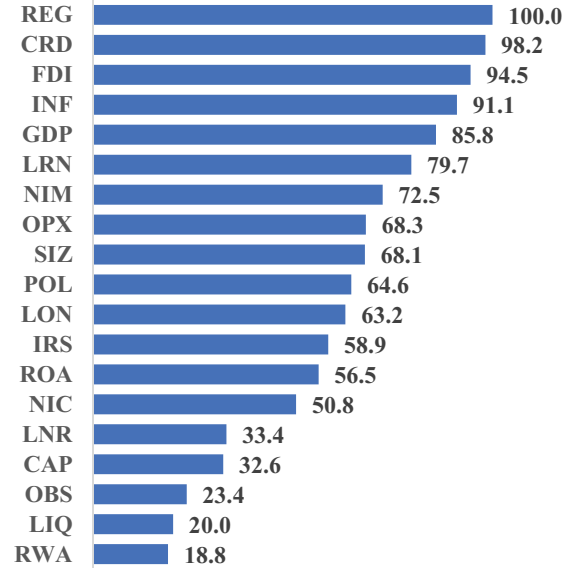


Figure 2. Relative Score
(Dependent variable: NIM).

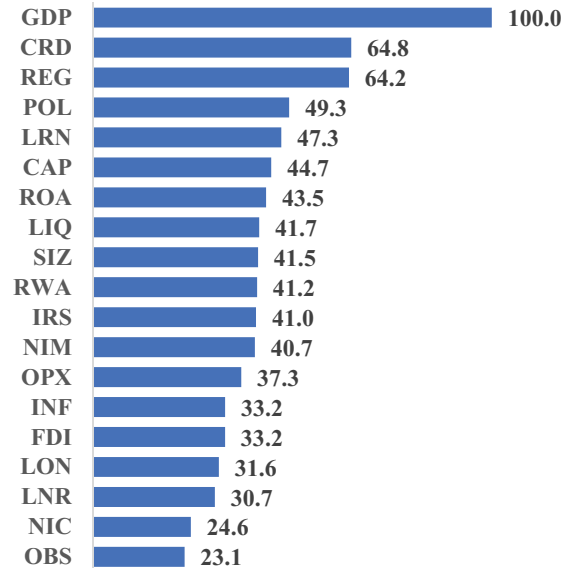


ability to obtain financing at attractive rates thus affecting its NIM. One interesting factor is non-interest income (NIC) which appears among the top ten determinants of NIM despite having no direct relation to interest rates. This observation is explained by the fact that banks rely on financing products to attract other non-interest related sources of income such as structuring, issuance, advisory and other fee-based products and vice versa, therefore creating a strong link between NIM and NIC.

**Figure 3. Relative Score
(Dependent variable: NPL).**

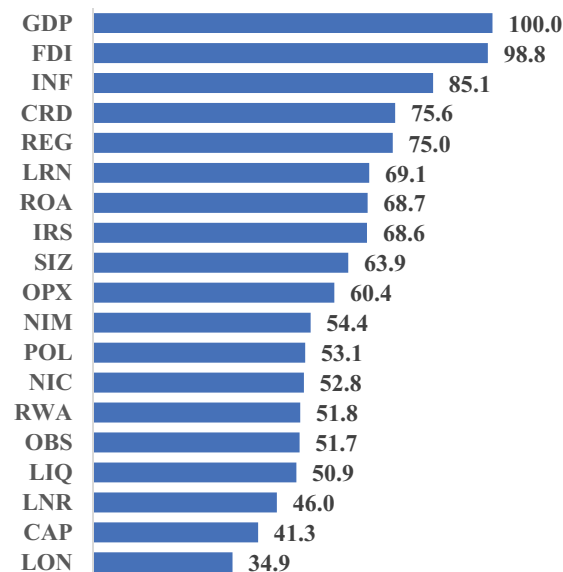


**Figure 4. Relative Score
(Dependent variable: ZSC).**



Figures 3 to 5 show the RVI values for the models containing our risk indicators (NPL, ZSC and SVL, respectively) as dependent variables. In contrast to the results of the models using the profitability measures, the current models show that country-level variables dominate in terms of their relative importance in determining the various risk indicators with at least four of the top five determinants being

Figure 5. Relative Score
(Dependent variable: SVL).



country-level variables. This suggests that the riskiness of the bank's portfolio is largely determined by macro rather than micro-related factors, especially those related to the growth of domestic credit and the business cycle as well as the development of the local financial markets and regulatory environment. Overall, it is worth noting that while models using profitability measures appear to have some dominant top determinants with twice the relative value of other determinants, the models using risk measures, appear to have more distributed RVI values with no significant domination by any one or two variables. There are several economic and policy implications arising from our findings in this study. First, the relatively high importance of country-level variables in determining the bank's riskiness suggest that regulators and policy makers have a big role to play in the process. It also suggests that steps to improve the stability of the banking system should start at the level of the regulations and the environment in which the banks conduct their business. Second, our results show that size and market power play an important positive role in impacting a bank's profitability and riskiness. This implies that smaller banks may have relatively lower profitability and riskier assets, which suggests the need for some policies and regulations to be specifically targeted towards ensuring the soundness and stability of this group of institutions. This is in contrast with the recent regulatory changes which focused primarily on regulating the systemically important banks. Lastly, the finding regarding the relatively important role played by operating efficiency in impacting the bank's profitability indicates the presence of some disparity in how efficiently banks operate. This suggests that significant improvements in profitability can be brought by improving operational efficiency in some banks.

6. Conclusion

We use the relative value importance indicator from the random forest model to analyse the comparative importance of a number of determinants of bank risk and profitability. The analysis shows that a bank's profitability is largely determined by bank-specific factors, while a bank's risk is mainly impacted by country-level factors. We find that market power and size proxies play a significant role in determining both the bank's profitability and its risk profile. The results also suggest a major role for a country's financial development status and regulatory quality in impacting the bank's riskiness. Lastly, the analysis shows that a bank's risk profile is determined by a number of variables with close relative importance levels, while

a bank's profitability is determined by small number of dominant variables with the remaining variables having much less importance. Future research can focus on studying changes in banks' profitability and riskiness arising from sudden or structural changes to some of the important variables identified in this study. Analysing such events will provide better insights into the role of these variables in influencing the bank's profitability and riskiness and the magnitude of this impact. This will also provide regulators and policy makers with better understanding of what areas to focus on in order to ensure the soundness and stability of the banking system and its participants.

Funding

The author received no direct funding for this research.

Author details

Nawaf Almaskati¹

E-mail: nawafmask@hotmail.com

ORCID ID: <http://orcid.org/0000-0001-8698-7560>

¹ University of Waikato, Hamilton, New Zealand.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Citation information

Cite this article as: The determinants of bank profitability and risk: A random forest approach, Nawaf Almaskati, *Cogent Economics & Finance* (2022), 10: 2021479.

Notes

1. We select the entities based on the industry classification (Banks) in the Thomson Reuters Business Classification (TRBC) field.
2. We exclude countries with less than three banks in our final sample.

References

- Angbazo, L. (1997). Commercial bank net interest margins, default risk, interest-rate risk, and off-balance sheet banking. *Journal of Banking & Finance*, 21(1), 55–87. [https://doi.org/10.1016/S0378-4266\(96\)00025-8](https://doi.org/10.1016/S0378-4266(96)00025-8)
- Ash, D.-K., & Huizinga, H. (1999). Determinants of commercial bank interest margins and profitability: Some international evidence. *The World Bank Economic Review*, 13(2), 379–408. <https://doi.org/10.1093/wber/13.2.379>
- Athanasoglou, P. P., Brissimis, S. N., & Delis, M. D. (2008). Bank-specific, industry-specific and macroeconomic determinants of bank profitability. *Journal of International Financial Markets, Institutions and Money*, 18(2), 121–136. <https://doi.org/10.1016/j.intfin.2006.07.001>
- Berger, A. N., & Udell, P. (2014). How does capital affect bank performance during financial crises? *Journal of Financial Economics*, 109(1), 146–176. <https://doi.org/10.1016/j.jfineco.2013.02.008>
- Berger, A. N., Udell, P., & Udell, P. (2009). Bank competition and financial stability. *Journal of Financial Services Research*, 35(2), 99–118. <https://doi.org/10.1007/s10693-008-0050-7>
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Bolt, W., de Haan, L., Hoeberichts, M., van Oordt, M. R. C., & Swank, J. (2012). Bank profitability during recessions. *Journal of Banking & Finance*, 36(9), 2552–2564. <https://doi.org/10.1016/j.jbankfin.2012.05.011>
- Bourke, P. (1989). Concentration and other determinants of bank profitability in Europe, North America and Australia. *Journal of Banking & Finance*, 13(1), 65–79. [https://doi.org/10.1016/0378-4266\(89\)90020-4](https://doi.org/10.1016/0378-4266(89)90020-4)
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Claeys, S., & Vander Vennet, R. (2008). Determinants of bank interest margins in Central and Eastern Europe: A comparison with the West. *Economic Systems*, 32(2), 197–216. <https://doi.org/10.1016/j.ecosys.2007.04.001>
- Dietrich, A., & Wanzenried, G. (2014). The determinants of commercial banking profitability in low-, middle-, and high-income countries. *The Quarterly Review of Economics and Finance*, 54(3), 337–354. <https://doi.org/10.1016/j.qref.2014.03.001>
- Friedman, J. H., & Meulman, J. J. (2003). Multiple additive regression trees with application in epidemiology. *Statistics in Medicine*, 22(9), 1365–1381. <https://doi.org/10.1002/sim.1501>
- Goddard, J., Molyneux, P., & Wilson, J. O. S. (2004). The profitability of European banks: A cross-sectional and dynamic panel analysis. *The Manchester School*, 72(3), 363–381. <https://doi.org/10.1111/j.1467-9957.2004.00397.x>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer Science & Business Media.
- Jones, S., Johnstone, D., & Wilson, R. (2017). Predicting corporate bankruptcy: An evaluation of alternative statistical frameworks. *Journal of Business Finance & Accounting*, 44(1–2), 3–34. <https://doi.org/10.1111/jbfa.12218>
- Li, X., Feng, H., Zhao, S., & Carter, D. A. (2021). The effect of revenue diversification on bank profitability and risk during the COVID-19 pandemic. *Finance Research Letters*, 101957.
- Maudos, J., & Solis, L. (2009). The determinants of net interest income in the Mexican banking system: An integrated model. *Journal of Banking & Finance*, 33(10), 1920–1931. <https://doi.org/10.1016/j.jbankfin.2009.04.012>
- Maudos, J. (2017). Income structure, profitability and risk in the European banking sector: The impact of the crisis. *Research in International Business and Finance*, 39, 85–101. <https://doi.org/10.1016/j.ribaf.2016.07.034>
- Mirzaei, A., Moore, T., & Liu, G. (2013). Does market structure matter on banks' profitability and stability? Emerging vs. advanced economies. *Journal of Banking & Finance*, 37(8), 2920–2937. <https://doi.org/10.1016/j.jbankfin.2013.04.031>
- Molyneux, P., & Thornton, J. (1992). Determinants of European bank profitability: A note. *Journal of Banking & Finance*, 16(6), 1173–1178. [https://doi.org/10.1016/0378-4266\(92\)90065-8](https://doi.org/10.1016/0378-4266(92)90065-8)
- Pasiouras, F., & Kosmidou, K. (2007). Factors influencing the profitability of domestic and foreign commercial banks

- in the European Union. *Research in International Business and Finance*, 21(2), 222–237. <https://doi.org/10.1016/j.ribaf.2006.03.007>
- Saunders, A., & Schumacher, L. (2000). The determinants of bank interest rate margins: An international study. *Journal of International Money and Finance*, 19(6), 813–832. [https://doi.org/10.1016/S0261-5606\(00\)00033-4](https://doi.org/10.1016/S0261-5606(00)00033-4)
- Short, B. K. (1979). The relation between commercial bank profit rates and banking concentration in Canada, Western Europe, and Japan. *Journal of Banking & Finance*, 3(3), 209–219. [https://doi.org/10.1016/0378-4266\(79\)90016-5](https://doi.org/10.1016/0378-4266(79)90016-5)
- Tregenna, F. (2009). The fat years: The structure and profitability of the US banking sector in the pre-crisis period. *Cambridge Journal of Economics*, 33(4), 609–632. <https://doi.org/10.1093/cje/bep025>
- Valverde, S. C., & Fernández, F. R. (2007). The determinants of bank margins in European banking. *Journal of Banking & Finance*, 31(7), 2043–2063. <https://doi.org/10.1016/j.jbankfin.2006.06.017>
- Yanikkaya, H., Gumus, N., & Pabuccu, Y. U. (2018). How profitability differs between conventional and Islamic banks: A dynamic panel data approach. *Pacific-Basin Finance Journal*, 48, 99–111. <https://doi.org/10.1016/j.pacfin.2018.01.006>



© 2022 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format.

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Economics & Finance (ISSN: 2332-2039) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

