What influences a bank's decision to go public?

by

Georgios Sermpinis, Serafeim Tsoukas, Ping Zhang
University of Glasgow[1]

February 2019

Abstract

A bank's decision to go public by issuing an Initial Public Offering (IPO) transforms its operations and capital structure. Much of the empirical investigation in this area focuses on the determinants of the IPO decision, applying accounting ratios and other publicly available information in non-linear models. We mark a break with this literature by offering methodological extensions as well as an extensive and updated US dataset to predict bank IPOs. Combining the least absolute shrinkage and selection operator (LASSO) with a cox proportional hazard, we uncover value in several financial factors as well as market-driven and macroeconomic variables, in predicting a bank's decision to go public. Importantly, we document a significant improvement in the model's predictive ability compared to standard frameworks used in the literature. Finally, we show that the sensitivity of a bank's IPO to financial characteristics is higher during periods of global financial crisis than in calmer times.

Key words: Equity financing, US banks, financial ratios, LASSO, forecasting

JEL: D40, G33, F23, C25, O16

1. Introduction

[1] We are grateful for comments and inputs from an anonymous referee. The third author is grateful to China Scholarship Council for funding (Grant number: 201508060273). Any remaining errors are our own. Correspondence to Georgios Sermpinis, email: Georgios.Sermpinis@glasgow.ac.uk.

Market finance in the USA has become an important source of funding for banks. According to the Federal Reserve Board, over the period 1996 to 2016 the net new issuance of US financial corporate equities outstanding more than tripled, from less than $50 billion to over $150 billion. The same body reports that the market value of total US corporate equity issues has risen from about $8 trillion in 1996 to around $36 trillion in 2016. This means that market participants have taken advantage of economic conditions as interest rates fell to historic lows. But not all banks were in a position to benefit from these unusual conditions as some financial institutions may rely less on equity financing as their funding largely comes from customer deposits and inter-bank financing.[2] However, during economic downsides banks face tighter lending conditions and increased levels of interbank borrowing costs (see Iyer et al, 2013 and Farinha et al, 2019). Using new estimation techniques across an extensive sample period that covers both periods of crisis and calmer times, the present study aims to identify the factors that influence a bank's desire to issue an equity IPO.

Our study considers the influence of bank-level financial information as well as market-level indicators; we ask how these explicators influence the decision at the level of the bank to issue stocks for the first time. The focus is on the decision of a bank to go public by issuing an Initial Public Offering (IPO), which is a financially significant step for a bank and provides new opportunities for financial flexibility, increased liquidity, better diversification, and attracting potential investors (Amihud and Mendelson, 1988; Pagano, 1993; Lowry, 2003; Bodnaruk et al., 2007; Kim & Weisbach, 2008 and Lowry et al., 2017). In addition, Houge and Loughran (1999) demonstrate that a bank's IPO decision can help managers to satisfy regulatory capital requirements, sell overvalued stock, and take advantage of better growth opportunities. After going public, Harris and Raviv (2014) indicate that the conditions of underlying market discipline and capital markets have more considerable influence on a public bank's ability to take risk than on a private bank. Samet et al. (2018) further clarify that public banks are able to take less credit risk during non-crisis periods compared to private banks. Moreover, if banks go public, market discipline can improve credibility and transparency in the banking industry and force public banks to maintain operational quality because of regular announcements of their financial health (Delis et al., 2011).

[2] For details on sources of external finance and firms' performance see Mallick and Yang (2011).

In this paper, we extend the literature methodologically, by developing a series of Cox proportional hazard, discrete hazard and logistic models combined with a more intuitive, yet innovative model, which is based on the variable selection technique, pioneered by Tibshirani (1996)—the least absolute shrinkage and selection operator (LASSO). This model, also known as L1 norm penalty, has proved very useful in identifying the most relevant predictors from an extensive set of candidate variables, without considering a pre-selection of these potential variables (van de Geer, 2008). The LASSO selection approach has a number of appealing characteristics: it not only helps identify the most relevant predictors from an extensive set of candidate variables, but it also improves the predictive power (Fan and Li, 2001 and Tian et al, 2015). In addition, LASSO does not require strict assumptions such as a pre-selection of the variables considered, and it is consistent statistically, as the number of observations approaches infinity (van de Geer, 2008). Importantly, LASSO can potentially sidestep the problem of multicollinearity, which is fairly common in reduced-form models, and it is computationally efficient even when considering a large set of potential predictors.

An additional important contribution of the present paper is that we test the estimator with superior predictive ability utilising a panel of US banks over an extensive time period. This approach not only allows us to compare our results with previous research, but also consider different time periods. Intuitively, banks respond in a different manner to extreme economic events as opposed to non-crisis periods, when they time their IPOs. Our sample covers the most recent global financial crisis as well as calmer (pre and post crisis) periods. We argue that across time periods, there is a differential sensitivity to bank and market information when it comes to the probability of banks going public.

To preview our findings, we discover value in several bank-specific financial factors as well as market-driven and macroeconomic variables in predicting the decision of banks to go public. In terms of the models' predictive ability, when we apply the LASSO estimator in a cox proportional hazard model, we note a significant improvement in predicting a bank's IPO and the penalized cox proportional hazard model outperforms other candidates. Specifically, we note improvements compared to a cox proportional hazard, discrete hazard and logistic models with or without LASSO. On the other hand, we show that the cox proportional hazard model underperforms discrete hazard and logistic models, which highlights the effect of LASSO on our algorithms. Our L1 penalized models are tuned through the AIC and the BIC criteria. We observe increased predictability on our dataset when the latter criterion is applied. Finally, when we put the model with superior predictive ability to the data and split our sample into crisis period and non-crisis periods, we find that the above variables become more potent in determining banks' IPOs, which signifies the ability of banks to time their IPOs relative to the economic conditions.

The rest of this work is laid out as follows. In section 2, we present a brief overview of the relevant literature. Sections 3 and 4 contain the data statistics and methodologies, respectively. Section 5 explains the empirical results of the forecasting simulation and Section 6 presents the econometric results of an empirical application. Section 7 provides conclusions.

2. Literature Review

Pagano et al. (1998) present the first systematic study on the determinants of firms' IPO. They rely on a number of accounting indicators to predict Italian firms' probability of issuing an IPO and conclude that larger firms, those with greater growth rates, or improved future investment opportunities, are more likely to go public. Several other studies confirm the importance of financial health in determining access to the public market in Germany (Boehmer and Ljungqvist, 2004), the UK and India (Albornoz and Pope, 2004 and Mayur and Kumar, 2016). In a slightly different setting, Helwege and Packer (2003) exploit the requirement of the Securities and Exchange Commission to obtain information about US public firms. The authors show that variables measuring size, profitability, leverage, interest coverage, R&D investment, capital structure, growth rate, future investment opportunities, ownership information and riskiness all have an important role in influencing the decision to issue an IPO.[3,4]

[3] This lends support to the finding of Pagano et al. (1998) and further demonstrates that issuing IPOs can be regarded as a primary mechanism to raise outside equity.

As well as evaluating the importance of financial information, Chemmanur et al. (2009) find that total factor productivity is a key contributor to the probability that a firm will issue IPOs. They find that a private firm which is larger, with a better growth rate and a higher total factor productivity is more likely to go public compared to its counterparts. Moreover, they show that if a private firm operates in an industry which is facing a higher degree of information asymmetry or costly evaluation of projects for outsiders, it is less likely to go public. Combined with the analysis of post-IPOs performance, the authors conclude that a firm is more likely to issue an IPO at the peak of its productivity cycle.

Taken from a different perspective, researchers have sought to explain the influence of the market environment on firms' IPO decisions. Subrahmanyam and Titman (1999) confirm that companies benefit more by issuing IPOs in a large, liquid public market. These benefits can encourage private companies to go public by issuing IPOs. Pástor and Veronesi (2005) further indicate that private firms are more likely to make IPO decisions when market conditions improve or stock prices increase. In contrast to this, Helwege and Liang (2004) indicate that the clusters of IPOs in stock markets are positively associated with investor optimism and are not related to the characteristics of industries such as profitability or growth opportunities.

While the literature on firms' IPOs is vast, the decision of banks to go public is less well studied. Ahmad and Kashian (2009) use a cox proportional hazard model and demonstrate that the quality of both assets and loans is linked to the IPO decision for credit unions that have converted into mutual savings institutions. On the other hand, the return on equity, the ratio of total loans to total assets and the size of the institution are not important determinants of a bank's decision to go public. Francis et al. (2009) use 272 US banks from the Securities Data Company (SDC) Global New Issues database in the logistical regression to distinguish the IPO decision from mergers and acquisitions. They demonstrate that a bank is less likely to go public during difficult economic times. In addition, Geyfman (2014) employs a cross-sectional dataset including 208 large commercial banks among 20 transition economies in Central and Eastern Europe and the former Commonwealth of Independent

States in 2010. The author finds that banks operating in advanced and mature markets are more inclined to go public, highlighting the role of financial architecture.

3. Data and summary statistics

3.1 Data description

Our dataset is drawn from the quarterly accounting reports taken from the Orbis Bank Focus database, published by Bureau Van Dijk Electronic Publishing (BvDEP). The Bank Focus database provides information on almost 40,000 institutions across the globe, with detailed coverage in the US over the period 1996–2016. We rely on Orbis Bank Focus to identify the banks' IPO date. The distribution of public and private banks studied is presented in Table 1.

Insert Table 1

Data on market indicators and macroeconomic variables are sourced from Bloomberg. These data items are reported quarterly. Following commonly used selection criteria in the literature, we exclude banks that do not have complete records on our explanatory variables and bank quarters with negative sales and assets. To control for the potential influence of outliers, we winsorize the regression variables at the 1st and 99th percentiles.

3.2 Choice of explanatory variables

Our models are supplied with forty-two potential explanatory variables, which can be divided into the following broad categories: bank-specific indicators, industry-specific predictors and macroeconomic variables. The choice of the explicators is based on a series of related studies (Pagano et al., 1998; Brau et al., 2003; Helwege & Packer, 2003; Pástor and Veronesi, 2005; Adjei et al., 2007; Ahmad and Kashian, 2009; Chemmanur et al., 2009; Tregenna, 2009; and Geyfman, 2014). To begin with the bank accounting variables, which measure various aspects of banks' health, these potential predictors are related to the determinants of CAMELS ratings. Specifically, they are aimed at assessing the overall safety and soundness of banks, covering capital adequacy; asset quality; management quality; earnings; liquidity; and sensitivity to market risk. Next, our industry-specific variables capture market concentration. Finally, we allow for fourteen macro-economic covariates that are likely to influence the timing of a bank's IPO.[5]

3.3 Summary statistics

We report summary statistics of the variables used in the empirical models in Table 2. We also present p-values for the tests of equality of means across the public and private banks in column 5 of Table 2. We observe, as expected, that public banks' size, growth rate, market share and income diversification are higher compared to private banks. On the other hand, capital, leverage and deposits in public banks are lower than in private banks. These statistics imply that public banks may absorb more growth opportunities from the stock market to enhance their performance and reduce risk. Overall, the tests point to significant differences between the two groups, which indicate that there is a correlation between banking activities and the decision about IPOs. Moving to the industry-specific indicators, we find significant differences between public and private banks, suggesting a link between the market climate and a bank's likelihood of going public.

Insert Table 2

4. Methodology

4.1 Cox proportional hazard model (CPH)

The CPH studies the effect of variables upon the time a specified event (IPO issuance) takes to happen. This analysis is often used to study problems that involve the passage of time before a certain event occurs. For example, studies on firm survival typically rely on hazard models to estimate the firm's chances of bankruptcy. In our context, the CPH model models

[5] For detailed definitions and abbreviations of all variables see Table A1 in the Appendix. Table A2 presents the cross-correlations between the bank-specific variables. It is generally observed that some variables exhibit relatively high correlation with each other, with some exceptions for variables that measure similar dimensions (e.g. banks' profitability using ROAA and ROAE). We note, however, that our preferred empirical methodology will carefully address this issue.

the likelihood of a bank issuing its equity IPO in a given (last) quarter, conditional on the fact that the bank did not undertake an IPO in any of the previous quarters. In other words, the hazard rate is the firm's probability of undertaking its bond IPO in a given quarter. Thus, the probability that a bank will issue an IPO takes the form:

$$Pr\left(Y_{i,t} = 1 \middle| Y_{i,t-1} = 0, X_{i,t}\right) = h\left(t, X_{i,t}\right)$$
$$= h_0(t, 0)\exp\left(\beta' X_{i,t}\right), \quad (1)$$

where $Y_{i,t}$ is equal to 1 if a bank issues an IPO in the public market and 0 otherwise; $h\left(t, X_{i,t}\right)$ is the hazard rate at time t for a bank controlling by a set $X_{i,t}$ of time-varying indicators including bank-specific, industry-specific and macroeconomic variables; β is a vector of unknown parameters to be estimated and $h_0(t, 0)$ is the baseline hazard function. The model is estimated by maximizing a partial-likelihood function.

4.2 Discrete hazard model (DH)

A bank can issue an IPO at any time within a quarter, but this event can only be observed when the information is released at the end of the corresponding quarter. The DH model is a discrete-time extension of the CPH that can capture this characteristic of our dataset, and its estimation can be applied by the complementary log-log (cloglog) model (Grilli, 2005; Jenkins 2005; Rabe-Hesketh and Skrondal, 2012). Our model takes the form:

$$cloglog(h(t)^D) = ln\{-ln(1 - h(t)^D)\}$$
$$= \gamma_0 + \gamma' X_{i,t}, \quad (2)$$

where $\gamma_0$is the baseline hazard rate; γ are estimated coefficient vectors and $X_{i,t}$ is our dataset.

4.3 Logistic model

Our dependent variable is binary and thus we also consider the logistic model, which is commonly used in the literature. The probability that a bank will issue an IPO based on the logistic model is:

$$Pr\left(Y_{i,t} = 1 \middle| X_{i,t}\right) = \frac{e^{\delta_0 + \delta' X_{i,t}}}{1 + e^{\delta_0 + \delta' X_{i,t}}}, \tag{3}$$

where δ_0 is the intercept to be estimated; δ are the estimated coefficient vectors in the logistical model and $X_{i,t}$ is our dataset.

4.4 LASSO

LASSO is a method of regression that enables estimation and variable selection simultaneously in a non-orthogonal setting (Tibshirani, 1996). Based on a shrinkage factor, LASSO selects variables by forcing some coefficients to zero and shrinking others. The variance of the estimated value is decreased while the accuracy of the regression prediction is increased.[6] Given a linear regression with standardized predictors and centred response values, LASSO resolves the l_1-penalized regression problem of estimating B to minimize:

$$\sum_{i=1}^{N} \left(Y_{i,t} - \mathrm{B}' X_{i,t}\right)^2, subject\ to \sum_{q=1}^{p} \left|\mathrm{B}_q\right| \leq s. \tag{4}$$

The above can be written in Lagrangian form as:

[6] See Sermpinis et al. (2018) for an application of LASSO estimators on firms' credit ratings.

$$\widehat{\mathrm{B}} = \arg min_{\mathrm{B}} \left(\sum_{i=1}^{N} \left(Y_{i,t} - \mathrm{B}' X_{i,t} \right)^2 + \lambda \sum_{q=1}^{p} \left| \mathrm{B}_q \right| \right). \qquad (5)$$

where $i = 1, 2, \ldots, N$ represents banks, $q = 1, 2 \ldots p$ indicates the surviving number of predictors with non-zero estimated coefficients and $t = 1, 2, \ldots, T$ represents different time periods. In equation (5), λ is the tuning parameter. The process of controlling different values of λ can be regarded as the procedure for selecting the number of independent variables in LASSO. As λ increases the sum of absolute values of estimated coefficients is reduced and shrinkage of coefficients is achieved. If λ exceeds a threshold value in the corresponding model, some estimated coefficients are ultimately set to zero. This procedure, the L1 norm penalty, generates a more interpretable and sparse model. Several approaches, such as cross-validation and information criteria, have been proposed in selecting the shrinkage factor λ. Zou et.al. (2007) provide an algorithm to obtain the optimal LASSO fit with the Akaike information criterion (AIC) (Akaike, 1974) and the Bayesian information criterion (BIC) (Schwartz, 1978)[7]. Sun and Zhang (2012) note that the computational cost of applying cross-validation in penalized models is considerable, while the theory of applying cross-validation is poorly understood. Therefore, the AIC and the BIC are used in selecting the tuning parameter λ in the LASSO and CPH, the DH and the logistic model combinations presented below.

As discussed above, LASSO provides more stable and restricted models (Tibshirani, 1996; Fan and Li, 2001). In addition, it is also a computationally simple and efficient method (Efron et al., 2004). Hence, these elements can lead to a superior predictability for its outputs (Tibshirani, 1996; Zou, 2006).

4.5 L1 Penalized Semi-Parametric Cox Proportional Hazard Model (Penalized CPH model)

[7] It is well-known that AIC and BIC have different properties in model selection (for details see Yang, 2005; Shao, 1997 and Zhang et al, 2010).

Tibshirani (1997) added the LASSO constraint form into the estimation of the CPH regression parameter and derived the L1 Penalized Semi-Parametric Cox Proportional Hazard Model. The LASSO estimator of the estimated coefficient β in the semi-parametric cox proportional hazard model is:

$$\hat{\beta} = argmax\ l(\beta)\ subject\ to\ \sum_{q=1}^{p} |\beta_q| \leq s, \tag{6}$$

where the likelihood function $l(\beta)$ is $l(\beta) = \sum_{Y_i\ uncensored} \left\{\beta^{'} X_{i,t} - \log\left(\sum_{Y_j \geq Y_i} \exp\left(\beta^{'} X_{j,t}\right)\right)\right\}$ in the semi-parametric cox proportional hazard model and $X_{i,t}$ contains the pool of the potential predictors in equation (1).

The above can be written into Lagrangian form as:

$$\hat{\beta} = argmin\{ -l(\beta) + \lambda \sum_{q=1}^{p} |\beta_q| \}. \tag{7}$$

In equation (7), as λ increases, the sum of absolute values of estimated coefficients is decreased and shrinkage of coefficients is achieved. If λ exceeds a threshold value in the corresponding models, some estimates are ultimately shrunk to zero. This ‘‘L1 norm penalty’’ generates a more interpretable and sparse cox model. All explanatory variables are standardized before applying the LASSO estimator.

4.6 L1 Penalized Discrete Hazard Model (Penalized DH model)

In the L1 Penalized Discrete Hazard Model, the LASSO parameter of the coefficient γ is estimated by maximizing the log-likelihood function with a L1-norm penalty placed on the sum of the absolute value of the covariate parameters. The model can be expressed as:

$$\hat{\gamma} = argmax\ l(\gamma)\ subject\ to\ \sum_{q=1}^{p} |\gamma_q| \leq s, \quad (8)$$

where the log-likelihood function $l(\gamma)$ is equal to $\sum_{i \in Q} w_{i,t} \ln\{F(\gamma_0 + \gamma' X_{i,t})\} + \sum_{i \notin Q} w_{i,t} \ln\{1 - F(\gamma_0 + \gamma' X_{i,t})\}$ where Q is the set of all observations that $Y_{i,t} = 1$ and $F(\gamma_0 + \gamma' X_{i,t}) = 1 - \exp\{-exp(\gamma_0 + \gamma' X_{i,t})\}$ and $w_{i,t}$ represents the optional weights in the discrete hazard model and $X_{i,t}$ is the same used in equation (2).

Or alternatively as:

$$\hat{\gamma} = argmin\{-l(\gamma) + \lambda \sum_{q=1}^{p} |\gamma_q|\}. \quad (9)$$

In line with the above-mentioned, all predictors are standardized before applying the LASSO estimator in this model.

4.7 L1 Penalized Logistic Model (Penalized Logistic Model)

The logistic model can be combined with LASSO as:

$$\hat{\delta} = argmaxl(\delta)\ subject\ to \sum_{q=1}^{p} |\delta_q| \leq s, \quad (10)$$

where $l(\delta)$ is $\sum\left(Y_{i,t}\left(\log\left(\frac{e^{\delta_0+\delta' X_{i,t}}}{1+e^{\delta_0+\delta' X_{i,t}}}\right)\right) + (1 - Y_{i,t})(\log(\frac{1}{1+e^{\delta_0+\delta' X_{i,t}}}))\right)$ in the corresponding logistic model and $X_{i,t}$ is the same in equation (3). All independent variables are standardized before implementing the LASSO estimator in logistic model.

The Lagrangian form is determined as:

$$\hat{\delta} = argmin\{-l(\delta) + \lambda \sum_{q=1}^{p} |\delta_q| \tag{11}$$

5. Predictive ability

We begin our analysis by presenting a forecasting simulation exercise to detect the model with the superior predictive ability. To measure the predictive performance of all competing models, we calculate the area under receiver operating characteristic curve (AUC), the accuracy ratio, and the brier score (see Duffie et al., 2007 and Tian et al., 2015). [8]

5.1 CPH and penalized CPH

Table 3 reports the AUC, the accuracy ratios and the brier scores for CPH and the penalized CPH. We benchmark their performance with a DH model, a logistic model and their two penalized variants. For the out-of-sample predictions of IPO decisions for banks, we use the past and current information and roll forward one step ahead of the prediction of the IPO decision. The initial estimation window is from 1996 to 2009.

Insert Table 3

In the in-sample, we note that the CPH presents an AUC of 24% and the penalized CPH, at 78%, is three times higher. LASSO seems to greatly improve the accuracy of the CPH model. The same trends can be observed from the accuracy ratios and the brier scores. In the out-of-sample, we note a similar improvement in terms of accuracy for the penalized CPH model compared to its simple CPH counterpart. The penalized CPH models are tuned, based on the

[8] For a detailed description of the tests please see section B in the Appendix.

AIC and BIC criteria. We note that the BIC models present slightly better accuracy in the out-of-sample. The BIC models also select a lower number of predictors compared to the AIC.

Concerning our benchmarks, we note that the DH model and the logistic model can provide more accurate in-sample and out-of-sample predictions than the simple CPH model. Adding the LASSO estimator in the DH model can improve the proportion of correct out-of-sample predictions from 45 percent to 53 percent. This increase in predictive performance by adding the LASSO estimator can also be noted in the logistic model, which confirms that predictive ability can be improved by adding LASSO. In general, we note that in the out-of-sample the penalized CPH model has the more accurate forecasts for the measures retained.

5.2 Predictive Deciles

To confirm the above-mentioned results, IPO decisions by out-of-sample prediction decile is reported in table 4. The decile method is frequently implemented in the default prediction (Shumway, 2001; Chava and Jarrow, 2004; Bharath and Shumway, 2008; Giordani et al., 2014; Tian et al., 2015; and Traczynski, 2017). The small changes in the predicted probabilities of IPO decisions do not have considerable influence on the decile in which a firm quarter lies in the distribution. The lowest probability of IPO decisions for banks would be included in the tenth decile and the highest would be in the first. Thus, the high proportion of banks appearing in the high probability for IPO decisions decile suggests high out-of-sample accuracy.

Insert Table 4

From table 4, we note that there is no observation in the first five percentiles in the CPH model, which suggests the lowest percentage of correct out-of-sample prediction among all candidate models (something that is consistent with the AUC values in the previous sections). The highest percentage of correct out-of-sample prediction is about 80%, which can be observed from the first five deciles in the penalized CPH model. This confirms the main conclusion from the above-mentioned evaluation methods (AUC, accuracy ratio and brier score) and demonstrates that the penalized CPH model outperforms all models studied in out-of-sample predictability.[9]

6. An empirical application using US data

Our findings thus far show that the penalized CPH model has substantial predictive ability compared to other models. We now present empirical evidence using data for US banks. Our extensive sample period covers the global financial crisis (2007-2009) which coincided with the collapse of the sub-prime mortgage lending market (Bekaert et al., 2014; Acharya & Mora, 2015; Dungey & Gajurel, 2015 and Ramcharan et al., 2016). We therefore have a unique opportunity to examine the sensitivity of our findings to different economic conditions. Motivated by this consideration, we split our sample into three parts: the pre-crisis period (1996-2006), the crisis period (2007-2009) and the post-crisis period (2010-2016). In the sub-sections below, we discuss our findings for each sub-sample separately. Tables 5 to 7 report the estimates of predictors employing the CPH model and its corresponding penalized versions[10] for each sub-period.[11] A positive coefficient indicates that an increase in that explanatory variable will improve the hazard of the IPO issue for a bank in any given quarter for a bank.

6.1 The pre-crisis period

To begin with the analysis of the CPH model, as shown in column 1 of table 5, we observe that most bank-specific determinants behave according to our expectations. Specifically, an increase in the bank's size (LNDETAS) reduces the hazard of the IPO issue in any given quarter and the estimate of LNDETAS2 illustrates a non-linear effect. These findings can be

---

[9] As a robustness test, we re-estimated our models using non-linear transformations of the bank-specific variables. The number of explanatory variables is 64 (23*2 bank-specific variables + 4 industry-specific predictors + 14 macroeconomic variables). The results, which are reported in the appendix, Tables A4 and A5 show that there is no significant increase in the predictive accuracy by including non-linear bank-specific variables in the candidate models. We continue to observe that in the out-of-sample predictions, the penalized CPH model provides the highest accuracy ratios compared with other models. Therefore, we conclude that our models are robust to controlling for non-linearities in the bank-specific variables.

[10] As an additional test, we replaced all macroeconomic determinants with time fixed effects. Our results remain unaffected in all models.

[11] We opt for estimated coefficients instead of hazard ratios, since the direction of effects is more important than their magnitude.

interpreted as follows. As banks grow in size they are less likely to issue IPOs, but once they attain a certain size threshold the probability of issuing an IPO is positively associated with the bank's size.[12] This finding is not only statistically significant, but also economically important. A unit increase in LNDETAS is associated with a reduction of 76% in the hazard of IPO issuance. As for banks' profitability (NETINTMAR), we find that it is negatively related to the probability of issuance. TIER1CAPTAS measures a bank's leverage and its estimated coefficient is positive and highly significant. A unit increase in this indicator (TIER1CAPTAS) improves the chances of an IPO issuance by 25%. This finding illustrates that banks with higher leverage are more likely to go public. The above findings on leverage and profitability suggest that banks with lower profitability and higher leverage are likely to make an IPO issue to diversify the credit risk (Albornoz and Pope, 2004 and Kim and Weisbach, 2008). Finally, a decrease in capital (TCAPTAS) is likely to increase the probability of a bank going public. This is linked with the preliminary and intuitive consideration of going public, which is to tap into different sources of capital (Lowry et al, 2017).

Insert Table 5

At the next stage, we add the penalized function into the CPH model. It should be noted that all surviving predictors after the penalty estimation are efficient variables that have predictive ability regarding the banks' decision to issue an IPO. Thus, p-values are calculated under post-selection after fitting the LASSO with a fixed value of tuning parameter, since the estimated coefficients are shrunk in LASSO estimation to select the "best" model. Compared to the CPH model, there exist 22 surviving explanatory variables in the L1 penalized CPH model with AIC-type and 16 in the BIC-type tuning parameter selector. All surviving determinants in the L1 penalized model contain bank-specific, industry-specific and macroeconomic factors.

In the L1 penalized CPH model with AIC-type tuning parameter selector, as shown in column 2, LNDETAS, GROAS and MASA are selected, and they are statistically significant. The sign of

[12] There is a line of thinking that argues the idea of "too big to fail" (TBTF) in the banking industry. Boyd and Heitz (2016) note that larger banks suffer more costs than benefits from TBTF in comparison with small and medium-sized banks. Therefore, larger banks may not go public because of the burden of the TBTF cost.

LNDETAS and MASA is negative, which is consistent with the findings in the baseline model (CPH). The estimate of GROAS shows that an increase in asset growth rate can lead to an increase improvement in the possibility of banks going public. This is in line with previous work which notes that banks with more investment opportunities are more likely to raise external finance (Pagano et al., 1998). PROGRO measures the productivity growth rate and its estimate demonstrates that a bank with a higher productivity growth rate is less likely to go public. This suggests that banks can operate efficiently using internal funds and therefore may be less inclined to source external finance. As for operating expenses, OPEXPTAS is positive and statistically significant, which implies that a bank with lower management efficiency is more likely to go public.

With respect to industry and macroeconomic indicators, HHI3 is the only industry-specific variable that is kept in the model and that is statistically significant under the post-selection. The negative estimate of this predictor illustrates the high degree of concentration in the banking industry that may prevent private banks going public. This confirms the findings of Grullon et al. (2017) that in industries with a relatively high concentration level in the US, firms can acquire more profits from mergers and acquisitions than from IPOs. Almost all macro-economic predictors are selected after the penalty. Overall, it appears that banks time their decision to go public and the probability of issuing is positively correlated with booming economic conditions.

In the L1 penalized CPH model with BIC-type tuning parameter selector, as reported in column 3, the selected predictors are slightly different from those in the model with the AIC-selector. In particular, no industry-related variables are included, while bank-specific variables such as liquidity, profitability, capital, leverage, operating expenses management and market share of a bank are found to be important determinants of the bank's IPO. Finally, several macroeconomic variables such as RSP500, CPI, LNGDPCAP and GRM1 are chosen in the model and are statistically significant under post-selection in line with the model that uses the AIC-type selector.

6.2 The crisis period

Starting with the analysis of the CPH model, as reported in column 1 of Table 6, most bank-specific variables are statistically significant. Importantly, the absolute value of these variables is higher compared to the pre-crisis period. This is a key finding which suggests that bank-specific variables are quantitatively more important predictors of IPOs during extreme economic conditions.[13] For example, the estimate of GROAS is only 0.0098 and not statistically significant in the pre-crisis period, while it increases to 0.1764 and becomes statistically significant at the 1% level in the crisis period.

Insert Table 6

Next, we find that in the L1 penalized CPH model with the AIC-type selector, LNDETAS, LOAAS, LIQASTAS, TCAPTAS, INCDIV and GRGDP are all statistically significant under post-selection from 13 surviving predictors in column 2. For the BIC-type selector counterpart, only four bank-specific variables are statistically significant among all seven selected variables, namely LNDETAS, LIQASTAS, DEPSTFUNTAS and INCDIV. We note that the number of variables that are statistically significant for our CPH models is higher than the one before and after the crisis sub-samples. These results signify the complexity of IPOs issuance under financial stress conditions. The number of factors that banks needs to consider in order to issue IPO is higher when a financial crisis is present. In other words, predicting a bank's decision to issue IPOs becomes a more strenuous task.[14]

6.3 The Post-crisis period

We now focus on the aftermath of the crisis. To begin with the analysis of the CPH model, as reported in column 1 of Table 7, most bank-specific variables enter with expected sign and retain their significance. However, the coefficients are significantly smaller compared to their counterparts in the crisis period. Regarding the industry-specific and macroeconomic variables, they are both statistically significant and the absolute value of the above-

[13] We report formal tests for the equality of coefficients across the sample periods in Table A.2 in the Appendix.

[14] Other research has connected the change in the IPO issuance behaviour before, during and after a financial crisis with changes in shareholder protection laws (Levine et. al., 2016) or with changes in the market structure (Weild and Kim, 2010).

mentioned variables in the post-crisis period is lower than in the crisis period. This indicates that while bank information and market conditions are important in affecting the probability of an IPO, they are less important than during the crisis period.

Insert Table 7

Moving to the analysis of penalized models, in the L1 penalized CPH model with the AIC-type selector, as shown in column 2, nine indicators of 18 surviving variables are statistically significant under post-selection. On the other hand, only three variables are statistically significant under post-selection from 8 selected variables in the penalized model with BIC-type selector. Comparing the magnitudes of the selected variables, we find, once again, that they are higher in the crisis period than in the post-crisis period.

7. Conclusion

The decision of a bank to go public by issuing an IPO is an important operational threshold event, which can lead to various investment and development plans for market participants. This paper uses quarterly data for US banks as original input in benchmark models and all competing models. We find that several bank-specific financial factors, market-driven and macroeconomic variables are important in predicting the decision of banks to go public. In terms of the models' predictive ability, when we apply the LASSO estimator in a cox proportional hazard model, we note a significant improvement in predicting a bank's IPO. The L1 penalized semi-parametric cox proportional hazard model provides the most accurate out-of-sample prediction among all candidate models. On the other hand, we show that the cox proportional hazard model underperforms discrete hazard and logistic models, which highlights the effect of LASSO on our algorithms. Our L1 penalized models are tuned through the AIC and the BIC criteria. We observe increased predictability on our dataset when the latter criterion is applied. Finally, when we split our sample into crisis and non-crisis periods, we find that bank-specific and macro variables become more potent in determining banks' IPOs, which signifies the ability of banks to time their IPOs relative to the economic conditions.

Our results should go forward in convincing market participants, policy makers and bankers about the utility of LASSO with a cox proportional hazard in predicting banks' IPOs. Our models manage to predict IPOs issuance accurately irrespective of the underlying financial environment. Our findings also reveal the change in structure of the relationship between IPOs and our dependent variables. This change is translated by the necessity of finance practitioners to be adaptive and follow the changes in the market environment. The number of variables and their significance in IPOs prediction is time-varying with the recent financial crisis causing a structural break in our estimations.

References

Acharya, V. and Mora, N. (2015). A Crisis of Banks as Liquidity Providers. *The Journal of Finance*, 70(1), 1-43.

Adjei, F., Cyree, K. and Walker, M. (2007). The Determinants and Survival of Reverse Mergers vs. IPOs. *Journal of Economics and Finance*, 32(2), 176-194.

Ahmad, Y. and Kashian, R. (2009). Modeling the Time to an Initial Public Offering: When Does the Fruit Ripen?. *Journal of Economics and Finance*, 34(4), 391-414.

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.

Albornoz, B. and Pope, P. (2004). *The Determinants of the Going Public Decision: Evidence from the U.K.* Mimeo

Allison, P. (1982). Discrete-Time Methods for the Analysis of Event Histories. *Sociological Methodology*, 13, 61.

Amihud, Y. and Mendelson, H. (1988). Liquidity and Asset Prices: Financial Management Implications. *Financial Management*, 17(1), 5.

Bekaert, G., Ehrmann, M., Fratzscher, M. and Mehl, A. (2014). The Global Crisis and Equity Market Contagion. *The Journal of Finance*, 69(6), 2597-2649.

Bharath, S. and Shumway, T. (2008). Forecasting Default with the Merton Distance to Default Model. *Review of Financial Studies*, 21(3), 1339-1369.

Bodnaruk, A., Kandel, E., Massa, M. and Simonov, A. (2007). Shareholder Diversification and the Decision to Go Public. *Review of Financial Studies*, 21(6), 2779-2824.

Boehmer, E. and Ljungqvist, A. (2004). On the Decision to Go Public: Evidence from Privately-held Firms. Mimeo

Boyd, J. and Heitz, A. (2016). The Social Costs and Benefits of Too-Big-To-Fail Banks: A "Bounding" Exercise. *Journal of Banking & Finance*, 68, 251-265.

Brau, J., Francis, B., and Kohers, N. (2003). The Choice of IPO versus Takeover: Empirical Evidence. *The Journal of Business*, 76(4), 583-612.

Chava, S. and Jarrow, R. (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance*, 8(4), 537-569.

Chemmanur, T., He, S. and Nandy, D. (2009). The Going-Public Decision and the Product Market. *Review of Financial Studies*, 23(5), 1855-1908.

Cox, D.R. (1972). Regression Models and Life Tables with Discussion. *Journal of the Royal Statistical Society: Series B*, 34, 187–220.

Delis, M., Molyneux, P. and Pasiouras, F. (2011). Regulations and Productivity Growth in Banking: Evidence from Transition Economies. *Journal of Money, Credit and Banking*, 43(4), 735-764.

Duffie, D., Saita, L. and Wang, K. (2007). Multi-Period Corporate Default Prediction with Stochastic Covariates. *Journal of Financial Economics*, 83(3), 635-665.

Dungey, M. and Gajurel, D. (2015). Contagion and Banking Crisis – International Evidence for 2007–2009. *Journal of Banking & Finance*, 60, 271-283.

Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004). Least Angle Regression. *The Annals of Statistics*, 32(2), 407–499.

Fan, J. and Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the American Statistical Association*, 96(456), 1348-1360.

Farinha, L., Spaliara, M.-E. and Tsoukas, S. (2019) Bank shocks and firm performance: New evidence from the sovereign debt crisis, forthcoming in *Journal of Financial Intermediation*.

Francis, B., Hasan, I. and Siregar, D. (2009). The Choice of IPO versus M&A: Evidence from the Banking Industry. *Applied Financial Economics*, 19(24), 1987-2007.

Geyfman, V. (2014). The Effect of Economic and Financial System Development on Banks' Listing Decisions: Evidence from Transition Economies. *Emerging Markets Finance and Trade*, 50(6), 174-192.

Giordani, P., Jacobson, T., Schedvin, E. and Villani, M. (2014). Taking the Twists into Account: Predicting Firm Bankruptcy Risk with Splines of Financial Ratios. *Journal of Financial and Quantitative Analysis*, 49(04), 1071-1099.

Grilli, L. (2005). The Random-Effects Proportional Hazards Model with Grouped Survival Data: A Comparison between the Grouped Continuous and Continuation Ratio Versions. *Journal of the Royal Statistical Society: Series A (Statistics In Society)*, 168(1), 83-94.

Grullon, G., Larkin, Y. and Michaely, R. (2017). *The Disappearance of Public Firms and the Changing Nature of U.S. Industries*. Mimeo

Harris, M. and Raviv, A. (2014). How to Get Banks to Take Less Risk and Disclose Bad News. *Journal of Financial Intermediation*, 23(4), 437-470.

Helwege, J. and Liang, N. (2004). Initial Public Offerings in Hot and Cold Markets. *Journal of Financial and Quantitative Analysis*, 39(03), 541.

Helwege, J. and Packer, F. (2003). *The Decision to Go Public: Evidence from Mandatory SEC Filings of Private Firms*. Mimeo

Hosmer Jr, D.W., Lemeshow, S. and Sturdivant, R.X. (2013). Applied Logistic Regression. 3rd ed. New Jersey: John Wiley & Sons.

Houge, T. and Loughran, T. (1999). Growth Fixation and the Performance of Bank Initial Public Offerings, 1983–1991. *Journal of Banking & Finance*, 23(8), 1277-1301.

Iyer, R., Peydro, J.-L., da Rocha-Lopes, S. and Schoar, A. (2013). Interbank liquidity crunch and the firm credit crunch: Evidence from the 2007–-2009 crisis, *Review of Financial Studies* 27, 347-372.

Jenkins, S.P. (2005). Survival Analysis. *Unpublished Manuscript, Institute for Social and Economic Research,* University of Essex, Colchester, UK.

Kim, W. and Weisbach, M. (2008). Motivations for Public Equity Offers: An International Perspective. *Journal of Financial Economics*, 87(2), 281-307.

Levine, R., Lin, C., and Xie, W. (2016). Spare tire? Stock markets, banking crises, and economic recoveries. *Journal of Financial Economics*, 120(1), 81-101.

Lowry, M. (2003). Why Does IPO Volume Fluctuate So Much? *Journal of Financial Economics*, 67(1), 3-40.

Lowry, M., Michaely, R. and Volkova, E. (2017). Initial Public Offerings: A Synthesis of the Literature and Directions for Future Research. *Foundations and Trends In Finance*, 11(3-4), 154-320.

Mallick, S. and Yang, Y. (2011). Sources of Financing, Profitability and Productivity: First Evidence from Matched Firms. *Financial Markets, Institutions and Instruments* 20, 221–252.

Mayur, M. and Kumar, M. (2016). *Determinants of Going-Public Decision in an Emerging Market: Evidence from India*. Mimeo

Pagano, M. (1993). The Flotation of Companies on the Stock Market. *European Economic Review*, 37(5), 1101-1125.

Pagano, M., Panetta, F. and Zingales, L. (1998). Why Do Companies Go Public? An Empirical Analysis. *The Journal of Finance*, 53(1), 27-64.

Pástor, Ľ. & Veronesi, P. (2005). Rational IPO Waves. *The Journal of Finance*, 60(4), 1713-1757.

Rabe-Hesketh, S. and Skrondal, A. (2012). *Multilevel and Longitudinal Modelling using Stata*. College Station, Tex.: Stata Press Publication.

Ramcharan, R., Verani, S. and Van Den Heuvel, S. (2016). From Wall Street to Main Street: The Impact of the Financial Crisis on Consumer Credit Supply. *The Journal of Finance*, 71(3), 1323-1356.

Samet, A., Boubakri, N. and Boubaker, S. (2018). Does Public–Private Status Affect Bank Risk Taking? Worldwide Evidence. *Journal of International Financial Markets, Institutions and Money*, 53, 287-306.

Sermpinis, G., Tsoukas, S. and Zhang, P. (2018). Modelling Market Implied Ratings Using LASSO Variable Selection Techniques. *Journal of Empirical Finance*, 48, 19-35.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461-464.

Shao, J. (1997). An Asymptotic Theory for Linear Model Selection. *Statistica Sinica.* 7(2), 221-242.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, 74(1), 101-124.

Subrahmanyam, A., and Titman, S. (1999). The Going-Public Decision and the Development of Financial Markets. *The Journal of Finance*, 54(3), 1045-1082.

Sun, T., and Zhang, C. (2012). Scaled Sparse Linear Regression. *Biometrika*, 99(4), 879-898.

Tian, S., Yu, Y. and Guo, H. (2015). Variable Selection and Corporate Bankruptcy Forecasts. *Journal of Banking & Finance*, 52, 89-100.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267-288.

Tibshirani, R. (1997). The Lasso Method for Variable Selection in the Cox Model. *Statistics in Medicine*, 16(4), 385-395.

Traczynski, J. (2017). Firm Default Prediction: A Bayesian Model-Averaging Approach. *Journal of Financial and Quantitative Analysis*, 52(03), 1211-1245.

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.

van de Geer, S. (2008). High-Dimensional Generalized Linear Models and the Lasso. *The Annals of Statistics*, 36(2), 614-645.

Yang, Y. (2005). Can the Strengths of AIC and BIC be Shared? A Conflict between Model Indentification and Regression Estimation. *Biometrika*, 92(4), 937-950.

Weild, D., and Kim, E. (2010). Market structure is causing the IPO crisis-and more. Capital Market Series, Grant Thornton.

Zhang, Y., Li, R. and Tsai, C. (2010). Regularization Parameter Selections via Generalized Information Criterion. *Journal of the American Statistical Association*, 105(489), 312-323.

Zou, H. (2006). The Adaptive Lasso and Its Oracle Properties. *Journal of the American Statistical Association*, 101(476), 1418–1429.

Zou, H., Hastie, T. and Tibshirani, R. (2007). On the “Degrees of Freedom” of the Lasso. *The Annals of Statistics.* 35, 2173.

Table 1 The distribution of banks

Year	Public Banks	Private Banks	Public banks percentage	Total
1996	176	345	33.78%	521
1997	194	406	32.33%	600
1998	203	467	30.30%	670
1999	222	603	26.91%	825
2000	229	677	25.28%	906
2001	242	742	24.59%	984
2002	265	6007	4.23%	6272
2003	276	6177	4.28%	6453
2004	285	6265	4.35%	6550
2005	301	6477	4.44%	6778
2006	251	5705	4.21%	5956
2007	256	5824	4.21%	6080
2008	237	5377	4.22%	5614
2009	257	5832	4.22%	6089
2010	275	5847	4.49%	6122
2011	283	5755	4.69%	6038
2012	301	6278	4.58%	6579
2013	307	6290	4.65%	6597
2014	319	6265	4.85%	6584
2015	298	5966	4.76%	6264
2016	292	5766	4.82%	6058

Notes: The table presents the distribution of banks by year.

Table 2 Summary statistics

Variable	Status	Mean	Standard Deviation	Minimum	Maximum	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Bank-specific						
LNDETAS	Public	14.334	1.475	10.896	19.800	0.000
	Private	12.120	1.188	9.328	16.760	
LNDETAS2	Public	24.063	2.949	17.187	34.996	0.000
	Private	19.635	2.375	14.052	28.915	
GROAS	Public	2.212	4.303	-7.052	34.731	0.000
	Private	1.553	4.037	-9.679	28.596	
LOAAS	Public	65.881	10.779	18.460	86.760	0.000
	Private	62.668	14.066	12.921	89.681	
EQAS	Public	9.965	2.349	4.376	28.822	0.000
	Private	10.766	3.154	4.473	52.197	
LIQASTAS	Public	5.336	3.890	1.018	35.805	0.000
	Private	9.039	7.126	1.055	51.237	
NETLOADEPSTFUN	Public	78.688	13.754	25.987	126.787	0.000
	Private	72.917	16.755	16.893	113.334	
NETLOATAS	Public	65.881	10.779	18.460	86.760	0.000
	Private	62.668	14.066	12.921	89.681	
DEPSTFUNTAS	Public	84.054	5.154	38.180	92.784	0.000
	Private	86.148	4.462	21.859	93.437	
LIQASDEPSTFUN	Public	6.341	4.663	1.291	50.953	0.000
	Private	10.474	8.235	1.350	60.950	
ROAA	Public	0.966	0.565	-5.249	2.890	0.012
	Private	0.950	0.741	-5.681	5.072	
ROAE	Public	10.011	6.114	-52.624	26.988	0.000
	Private	9.309	7.562	-52.168	35.841	
NETINTMAR	Public	3.960	0.762	1.500	7.350	0.000
	Private	4.057	0.825	1.381	7.999	
TCAPTAS	Public	10.354	2.059	5.432	29.297	0.000
	Private	11.101	3.035	5.526	50.690	
TIER1CAPTAS	Public	9.286	2.080	4.311	28.643	0.000
	Private	10.314	3.063	4.643	50.146	
LOALOSPROLOA	Public	0.108	0.155	-0.129	1.442	0.000
	Private	0.090	0.159	-0.144	1.671	
PROGRO	Public	0.720	9.249	-44.202	64.424	0.369
	Private	0.647	10.513	-49.428	66.672	
OPEXPTAS	Public	0.754	0.209	0.313	2.256	0.211
	Private	0.755	0.267	0.270	6.032	
COSINC	Public	64.386	12.073	35.777	158.992	0.000
	Private	68.609	15.566	31.229	178.738	
OVHTAS	Public	0.754	0.209	0.313	2.256	0.211
	Private	0.755	0.267	0.270	6.032	
MSAS	Public	0.112	0.415	0.001	6.978	0.000
	Private	0.004	0.011	0.000	0.249	
DEPLOA	Public	127.984	29.081	71.429	372.021	0.000
	Private	144.767	46.532	80.322	450.640	
DEPLOAGRO	Public	-0.108	3.774	-13.184	16.239	0.000
	Private	0.179	5.210	-17.411	21.725	
INCDIV	Public	22.592	11.542	-2.613	81.469	0.000
	Private	16.400	10.111	-5.800	95.082	
Industry-specific						
HHI3	Public	8.439	2.124	4.610	12.368	0.000
	Private	3.153	1.579	1.739	94.565	

HHI5	Public	9.415	1.827	6.209	12.972	0.000
	Private	3.731	1.584	2.292	94.568	
CON3	Public	48.917	6.158	36.667	60.279	0.000
	Private	29.781	4.725	22.743	98.826	
CON5	Public	62.483	4.487	54.539	70.307	0.000
	Private	39.698	4.855	32.906	99.603	

Notes: The Table reports summary statistics of the explanatory variables used in the empirical models. Column 5 reports the p-value for the test of equality of means between the public and private group. A detailed description of the variables used in this study is given in Table A1 in Appendix.

Table 3 Accuracy ratios and the number of surviving variables in the CPH model and its penalized versions

Model		CPH model	Penalized CPH model		DH model	Penalized DH model		Logistic model	Penalized Logistic model	
			AIC	BIC		AIC	BIC		AIC	BIC
AUC	In-sample	0.238	0.779	0.779	0.746	0.745	0.745	0.749	0.748	0.689
	Out-of-sample	0.210	0.793	0.797	0.450	0.533	0.533	0.466	0.455	0.577
AR	In-sample	-0.524	0.557	0.557	0.492	0.489	0.489	0.497	0.497	0.379
	Out-of-sample	-0.581	0.586	0.594	-0.101	0.067	0.067	-0.068	-0.090	0.153
BS	In-sample	0.756	0.112	0.112	0.101	0.105	0.105	0.101	0.102	0.111
	Out-of-sample	0.707	0.109	0.109	0.205	0.195	0.195	0.203	0.202	0.182
Surviving variables		42	23	21	42	35	5	42	39	8

Notes: CPH model represents the Cox proportional hazard model and DH model refers to the discrete hazard model. "AUC" refers to the area under receiver operating characteristic curve. "AR" stands for accuracy ratio. "BS" represents the brier score. "AIC" is the AIC-type tuning parameter selector. "BIC" is the BIC-type tuning parameter selector.

Table 4 IPO decision by out-of-sample prediction decile

Decile	CPH model	Penalized CPH model		DH model	Penalized DH model		Logistic model	Penalized Logistic model	
		AIC	BIC		AIC	BIC		AIC	BIC
1	0	64.22%	64.22%	11.01%	15.60%	15.60%	11.01%	13.76%	17.43%
2	0	4.59%	4.59%	5.50%	11.01%	11.01%	5.50%	4.59%	9.17%
3	0	4.59%	4.59%	10.09%	10.09%	10.09%	11.01%	11.93%	14.68%
4	0	5.50%	6.42%	8.26%	8.26%	8.26%	8.26%	7.34%	8.26%
5	0	2.75%	1.83%	8.26%	6.42%	6.42%	11.01%	7.34%	7.34%
6-10	100%	18.35%	18.34%	56.88%	48.62%	48.62%	53.21%	55.03%	43.13%
AUC	0.210	0.793	0.797	0.450	0.533	0.533	0.466	0.455	0.577

Notes: CPH model represents the Cox proportional hazard model. DH model refers to the discrete hazard model. "AUC" refers to the area under receiver operating characteristic curve. "AR" stands for accuracy ratio. "BS" represents the brier score. "AIC" is the AIC-type tuning parameter selector. "BIC" is the BIC-type tuning parameter selector.

Table 5 The estimates of candidate models in the pre-crisis period

Variable	CPH model	Penalized CPH model_AIC	Penalized CPH model_BIC
	(1)	(2)	(3)
LNDETAS	-1.4340***	-2.607E-07***	
	(0.0000)	(0.0000)	
LNDETAS2	0.7176***		
	(0.0000)		
GROAS	0.0098	0.0111***	
	(0.2758)	(0.0000)	
LOAAS	-0.0434	-1.129E-06***	
	(0.4133)	(0.0000)	
LIQASTAS	-0.0220		
	(0.7749)		
NETLOADEPSTFUN	0.0373	-0.0039	-0.0035***
	(0.3979)	(1.0000)	(0.0000)
NETLOATAS	NA		
DEPSTFUNTAS	0.0356	-0.0135	-0.0135***
	(0.4663)	(1.0000)	(0.0000)
LIQASDEPSTFUN	0.0292		
	(0.6115)		
ROAA	0.3717		
	(0.3247)		
ROAE	0.0055		
	(0.8340)		
NETINTMAR	-0.3862***	-0.1439	-0.1567
	(0.0141)	(1.0000)	(1.0000)
TCAPTAS	-0.2128**		-0.0330***
	(0.0499)		(0.0000)
EQAS	0.0665		
	(0.2318)		
TIER1CAPTAS	0.2208**	0.0480	0.0840***
	(0.0490)	(1.0000)	(0.0000)
LOALOSPROLOA	0.4510	-0.2152***	
	(0.5053)	(0.0000)	
PROGRO	-0.0056	-0.0028***	
	(0.3047)	(0.0000)	
OPEXPTAS	1.2440	0.4969***	0.4868***
	(0.1140)	(0.0000)	(0.0000)
COSINC	0.0037		
	(0.7937)		
OVHTAS	NA	0.0300	0.0386
		(1.0000)	(1.0000)
MSAS	-1.4630**	-0.9438***	-0.9142***
	(0.0338)	(0.0000)	(0.0000)
DEPLOA	0.0009		
	(0.7583)		
DEPLOAGRO	0.0041		
	(0.7011)		
INCDIV	-0.0133		
	(0.2292)		
HHI3	3.8430	-0.0246***	
	(0.5143)	(0.0000)	
HHI5	-4.0980		
	(0.4938)		
CON3	-0.4539		

	(0.4614)		
CON5	0.5354		
	(0.4179)		
RSP500	0.0034	0.0085***	0.0074***
	(0.7727)	(0.0000)	(0.0000)
CPI	0.1336	0.2122***	0.1632***
	(0.5310)	(0.0000)	(0.0000)
GRGDP	0.1943	0.0456***	0.0580
	(0.2679)	(0.0000)	(1.0000)
LNGDPCAP	-0.4085	-1.8429***	-2.2406***
	(0.8069)	(0.0000)	(0.0000)
GRGNP	-0.5861	-0.2704***	-0.3669
	(0.3543)	(0.0000)	(1.0000)
INTR_10Y	1.1540		
	(0.4249)		
INTR_10Y2	-11.6200		
	(0.3525)		
INTR_3M	-0.0796		
	(0.8650)		
INTR_3M2	4.4950		
	(0.3156)		
SLYC	NA	-0.1343***	
		(0.0000)	
SLYC2	3.0640	3.4246***	
	(0.6633)	(0.0000)	
GRM1	0.0239		-0.0169***
	(0.7645)		(0.0000)
GRM2	-0.1683	-0.1844***	-0.1637
	(0.3716)	(0.0000)	(1.0000)
HPI	0.1482	0.0656	0.0926
	(0.1351)	(1.0000)	(1.0000)

Notes: CPH model represents the Cox proportional hazard model. "AIC" is the AIC-type tuning parameter selector. "BIC" is the BIC-type tuning parameter selector. P-values related to z-statistics reported in the parentheses are Huber–White robust estimates, clustered at the firm level. *** denotes significance at the 1% level. ** denotes significance at the 5% level. * denotes significance at the 10% level.

Table 6 The estimates of candidate models in the crisis period

Variable	CPH model	Penalized CPH model_AIC	Penalized CPH model_BIC
	(1)	(2)	(3)
LNDETAS	-1.737E+03***	-0.0062*	-0.0072*
	(0.0000)	(0.0940)	(0.0560)
LNDETAS2	NA		
GROAS	0.1764***	-0.0244	
	(0.0000)	(0.9060)	
LOAAS	-1.1290***	-0.0191*	
	(0.0000)	(0.0980)	
LIQASTAS	-31.0100***	-0.1885*	-0.1905**
	(0.0000)	(0.0940)	(0.0390)
NETLOADEPSTFUN	1.1890***	-0.0016	
	(0.0000)	(0.9030)	
NETLOATAS	NA	-0.0008	
		(0.8690)	
DEPSTFUNTAS	1.6790***	-0.0715	-0.1137*
	(0.0000)	(0.9020)	(0.0530)
LIQASDEPSTFUN	25.1600***		
	(0.0000)		
ROAA	-21.1700***		
	(0.0000)		
ROAE	1.9640***		
	(0.0000)		
NETINTMAR	3.6150***		
	(0.0000)		
TCAPTAS	0.2698**	0.1632*	
	(0.0228)	(0.0990)	
EQAS	1.9360***		
	(0.0000)		
TIER1CAPTAS	0.1859		
	(0.1663)		
LOALOSPROLOA	-5.6620***		
	(0.0000)		
PROGRO	-0.1182***	0.0143	
	(0.0000)	(0.6290)	
OPEXPTAS	19.7500***		
	(0.0000)		
COSINC	0.0707		
	(0.2069)		
OVHTAS	NA		
MSAS	11.5500***		
	(0.0000)		
DEPLOA	NA		
DEPLOAGRO	-0.1145***		
	(0.0000)		
INCDIV	-0.1053**	0.0426*	0.0556*
	(0.0188)	(0.0970)	(0.0590)
HHI3	1.306E+05***		
	(0.0000)		
HHI5	NA		
CON3	NA		

CON5	NA		-0.0171 (0.9390)
RSP500	5.011E+04*** (0.0000)		
CPI	-1.589E+04*** (0.0000)	-0.4462 (0.1530)	
GRGDP	-2.026E+05*** (0.0000)	0.4019* (0.0920)	
LNGDPCAP	NA		
GRGNP	NA		3.1577 (0.1270)
INTR_10Y	NA		
INTR_10Y2	-2.9830 (0.6816)		
INTR_3M	NA		
INTR_3M2	NA		
SLYC	NA		
SLYC2	NA	-19.6895 (0.1870)	-31.4148 (0.7370)
GRM1	-2.388E+04*** (0.0000)		
GRM2	-1.916E+04*** (0.0000)		
HPI	NA		

Notes: CPH model represents the Cox proportional hazard model. “AIC” is the AIC-type tuning parameter selector. “BIC” is the BIC-type tuning parameter selector. P-values related to z-statistics reported in the parentheses are Huber–White robust estimates, clustered at the firm level. *** denotes significance at the 1% level. ** denotes significance at the 5% level. * denotes significance at the 10% level.

Table 7 The estimates of candidate models in the post-crisis period

Variable	CPH model	Penalized CPH model_AIC	Penalized CPH model_BIC
	(1)	(2)	(3)
LNDETAS	6.3560***		-0.0007
	(0.0002)		(0.3240)
LNDETAS2	-3.1730***		
	(0.0002)		
GROAS	0.0002		
	(0.9897)		
LOAAS	-0.3259***	-0.0143***	
	(0.0065)	(0.0000)	
LIQASTAS	-2.3900***		
	(0.0000)		
NETLOADEPSTFUN	0.2474***	-1.062E-07	
	(0.0078)	(1.0000)	
NETLOATAS	NA		
DEPSTFUNTAS	0.4191***	-0.0279	-0.0369*
	(0.0002)	(1.0000)	(0.0640)
LIQASDEPSTFUN	2.1610***	0.0229	
	(0.0000)	(1.0000)	
ROAA	1.5470***		
	(0.0002)		
ROAE	-0.2509***		
	(0.0000)		
NETINTMAR	0.8035**		
	(0.0367)		
TCAPTAS	0.9610***	0.0325***	
	(0.0000)	(0.0000)	
EQAS	-0.3103***	-0.0078	
	(0.0026)	(1.0000)	
TIER1CAPTAS	-0.6686***		
	(0.0003)		
LOALOSPROLOA	-2.8330***		
	(0.0005)		
PROGRO	-0.0314***	-0.0120***	
	(0.0000)	(0.0000)	
OPEXPTAS	-2.1830		
	(0.2147)		
COSINC	0.0125		
	(0.4718)		
OVHTAS	NA		
MSAS	-6.2670***	-0.4111	
	(0.0000)	(1.0000)	
DEPLOA	-0.0193*	-0.0095***	-0.0023
	(0.0755)	(0.0000)	(0.5630)
DEPLOAGRO	-0.0525**		
	(0.0147)		
INCDIV	0.0497***		
	(0.0094)		
HHI3	19.9700		
	(0.4986)		
HHI5	74.1200***		
	(0.0001)		
CON3	-10.7400**		-0.0442

	(0.0484)		(0.5500)
CON5	-17.7200***	-0.1094***	-0.0600
	(0.0000)	(0.0000)	(0.4740)
RSP500	0.3170***		
	(0.0008)		
CPI	0.3504	-0.1486***	
	(0.2414)	(0.0010)	
GRGDP	-4.7710***		
	(0.0008)		
LNGDPCAP	-78.1600**	-0.7421***	-0.8477
	(0.0235)	(0.0000)	(0.7980)
GRGNP	19.6100***		
	(0.0008)		
INTR_10Y	0.0011		0.0950**
	(0.9998)		(0.0450)
INTR_10Y2	-514.5000		
	(0.1355)		
INTR_3M	-26.4500***	-0.3214	
	(0.0092)	(0.0000)	
INTR_3M2	3.091E+03***	49.3428***	
	(0.0002)	(0.0000)	
SLYC	NA		
SLYC2	465.0000	3.1383	
	(0.1384)	(1.0000)	
GRM1	-1.4570***	-0.0738***	
	(0.0002)	(0.0000)	
GRM2	2.3200***	-0.0242	-0.1523*
	(0.0016)	(1.0000)	(0.0910)
HPI	-1.1830**	0.0120	
	(0.0359)	(0.9960)	

Notes: CPH model represents the Cox proportional hazard model. “AIC” is the AIC-type tuning parameter selector. “BIC” is the BIC-type tuning parameter selector. P-values related to z-statistics reported in the parentheses are Huber–White robust estimates, clustered at the firm level. *** denotes significance at the 1% level. ** denotes significance at the 5% level. * denotes significance at the 10% level.

Appendix

Section A

Table A.1 Variables definition and expected relationship

Variable	Definition	Label	Expected sign
Bank-specific (24)			
Size	The logarithm of total real assets	LNDETAS	+
	The logarithm of the square of real total assets	LNDETAS2	~
	The rate of growth of real assets	GROAS	+
Liquidity	Loans / assets	LOAAS	+
	Liquid asset / total assets	LIQASTAS	-
	Net loans / deposits and short-term funding	NETLOADEPSTFUN	+
	Net loans / total assets	NETLOATAS	+
	Deposits and short-term funding / total assets	DEPSTFUNTAS	-
	Liquid assets / deposits & short-term funding	LIQASDEPSTFUN	-
Profitability	The average return on equity	ROAA	-
	The average return on assets	ROAE	-
	Net interest margin	NETINTMAR	-
Capital	Capital to assets ratio	TCAPTAS	~
	Equity / assets	EQAS	~
Leverage	Tier 1 ratio	TIER1CAPTAS	+
Credit risk	Loan loss provisions / loans	LOALOSPROLOA	+
Productivity growth	Rate of change in inflation-adjusted gross total revenue / the number of employees	PROGRO	-
Operating expenses management	Operating expenses / total assets	OPEXPTAS	+
	Operating costs / Operating income ratio	COSINC	+
	Overheads to total assets	OVHTAS	+
Market share	Market share (in terms of assets) of individual banks	MSAS	~
Deposit	Total deposits / total loans	DEPLOA	~
	The growth rate of deposits	DEPLOAGRO	~
Income diversification	Non-interest income to total operating revenue	INCDIV	-
Industry-specific (4)			
Concentration	The three-firm Herfindahl-Hirschman index	HHI3	-
	The five-firm Herfindahl-Hirschman index	HHI5	-
	The assets of the three largest banks / the assets of all banks in the same dataset	CON3	-
	The assets of the five largest banks / the assets of all banks in the same dataset	CON5	-
Macroeconomic (14)			
Stock market performance	The return of S&P500	RSP500	+

Inflation rate	Current period inflation	CPI	~
GDP growth rate	The real gross domestic product (GDP) growth rate	GRGDP	~
GDP per capita	The logarithm of GDP per capita	LNGDPCAP	~
GNP growth rate	The GNP growth rate	GRGNP	~
Interest rate	10-year government bond yield	INTR_10Y	~
	The square of 10-year government bond yield	INTR_10Y2	~
	3-month interbank rate	INTR_3M	~
	The square of 3-month interbank rate	INTR_3M2	~
Slope of the yield curve	The difference between the 10-year government bond yield and the three-month interbank rate	SLYC	~
	The square of the abovementioned yield curve	SLYC2	~
Market growth	The growth rate in money supply (M1)	GRM1	~
	The growth rate in money supply (M2)	GRM2	~
House price growth rate	All-Transactions House Price Index for the United States	HPI	+

Notes: “+” indicates that the probability of a bank going public would improve if the covariates rise. “-” indicates that the probability of a bank going public would reduce if the covariates rise. “~” indicates uncertainty in the sign.

Table A.2 Cross-Correlations

Variable	LNDETAS	LNDETAS2	GROAS	LOAAS	LIQASTAS	NETLOADEPSTFUN	NETLOATAS	DEPSTFUNTAS	LIQASDEPSTFUN	ROAA	ROAE	NETINTMAR	TCAPTAS	EQAS	TIER1CAPTAS	LOALOSPROLOA	PROGRO	OPEXPTAS	COSINC	OVHTAS	MSAS	DEPLOA	DEPLOAGRO	INCDIV
LNDETAS	1.000																							
LNDETAS2	1.000	1.000																						
GROAS	-0.059	-0.059	1.000																					
LOAAS	-0.163	-0.163	0.034	1.000																				
LIQASTAS	0.066	0.066	0.159	-0.188	1.000																			
NETLOADEPSTFUN	-0.035	-0.035	0.013	0.924	-0.198	1.000																		
NETLOATAS	-0.163	-0.163	0.034	1.000	-0.188	0.924	1.000																	
DEPSTFUNTAS	-0.340	-0.340	0.048	-0.012	0.060	-0.385	-0.012	1.000																
LIQASDEPSTFUN	0.090	0.090	0.151	-0.195	0.985	-0.158	-0.195	-0.059	1.000															
ROAA	0.035	0.035	-0.108	-0.130	-0.091	-0.144	-0.130	0.074	-0.097	1.000														
ROAE	0.042	0.042	-0.042	-0.061	-0.066	-0.112	-0.061	0.163	-0.080	0.869	1.000													
NETINTMAR	-0.345	-0.345	-0.002	0.030	-0.112	-0.046	0.030	0.229	-0.125	0.421	0.308	1.000												
TCAPTAS	-0.113	-0.113	-0.054	-0.049	0.046	0.071	-0.049	-0.348	0.080	0.000	-0.260	0.057	1.000											
EQAS	0.067	0.067	-0.045	-0.174	0.017	-0.023	-0.174	-0.403	0.057	0.010	-0.310	0.093	0.830	1.000										
TIER1CAPTAS	-0.224	-0.224	-0.041	-0.086	0.039	0.020	-0.086	-0.298	0.069	0.004	-0.258	0.091	0.974	0.833	1.000									

LOALOSPROLOA	0.128	0.128	-0.018	0.066	-0.001	0.100	0.066	-0.117	0.007	-0.520	-0.517	-0.096	0.044	0.007	0.000	1.000								
PROGRO	0.002	0.002	0.069	0.016	0.015	0.018	0.016	-0.013	0.016	0.028	0.039	0.014	0.000	0.018	-0.002	0.065	1.000							
OPEXPTAS	-0.018	-0.018	-0.023	-0.024	0.147	-0.010	-0.024	-0.008	0.174	-0.115	-0.086	0.345	-0.006	-0.029	-0.028	0.049	0.018	1.000						
COSINC	-0.111	-0.111	0.121	0.021	0.161	0.022	0.021	-0.007	0.171	-0.740	-0.625	-0.190	0.028	0.013	0.033	0.119	-0.035	0.560	1.000					
OVHTAS	-0.018	-0.018	-0.023	-0.024	0.147	-0.010	-0.024	-0.008	0.174	-0.115	-0.086	0.345	-0.006	-0.029	-0.028	0.049	0.018	1.000	0.560	1.000				
MSAS	0.728	0.728	-0.037	-0.126	0.056	-0.036	-0.126	-0.234	0.074	0.098	0.092	-0.198	-0.093	0.042	-0.191	0.054	0.032	0.048	-0.099	0.048	1.000			
DEPLOA	-0.004	-0.004	0.013	-0.879	0.204	-0.924	-0.879	0.310	0.175	0.089	0.062	0.030	-0.042	0.042	0.012	-0.071	-0.008	0.004	0.020	0.004	-0.023	1.000		
DEPLOAGRO	0.031	0.031	0.310	-0.024	0.203	-0.023	-0.024	0.006	0.195	-0.044	-0.036	-0.050	-0.039	-0.056	-0.043	0.110	-0.016	-0.037	-0.020	-0.037	-0.009	0.046	1.000	
INCDIV	0.507	0.507	-0.059	-0.128	0.258	-0.039	-0.128	-0.223	0.288	0.105	0.146	-0.263	-0.091	-0.102	-0.172	0.066	0.019	0.519	0.094	0.519	0.444	0.003	0.039	1.000

Note: All bank-specific variables are as defined in Table A.1. The number in each cell indicates the correlation between the row and column variables.

Table A.3 Tests of equality of estimated coefficients

	CPH in the pre-crisis period	CPH in the crisis period	CPH in the post-crisis period
	(1)	(2)	(3)
LNDETAS	0.0000	0.0000	0.0312
LNDETAS2	-	-	-
GROAS	0.0000	0.0000	0.0000
LOAAS	0.2648	0.1121	0.4351
LIQASTAS	0.0000	0.0000	0.0000
NETLOADEPSTFUN	0.2005	0.2395	0.0842
NETLOATAS	-	-	-
DEPSTFUNTAS	0.0004	0.0001	0.0230
LIQASDEPSTFUN	0.0000	0.0000	0.0000
ROAA	0.0039	0.0009	0.0011
ROAE	0.0000	0.0000	0.0000
NETINTMAR	0.0319	0.0735	0.4098
TCAPTAS	0.0000	0.0030	0.0000
EQAS	0.0002	0.0008	0.0000
TIER1CAPTAS	0.0000	0.0000	0.0000
LOALOSPROLOA	0.0483	0.0654	0.0215
PROGRO	0.0000	0.0000	0.0000
OPEXPTAS	0.2972	0.7541	0.3845
COSINC	0.6639	0.3804	0.3694
OVHTAS	-	-	-
MSAS	0.1303	0.1434	0.0479
DEPLOA	-	-	-
DEPLOAGRO	0.1029	0.0336	0.0461
INCDIV	0.2599	0.8992	0.3605
HHI3	0.0000	0.0000	0.0000
HHI5	-	-	-
CON3	-	-	-
CON5	-	-	-
RSP500	0.0016	0.0332	0.0006
CPI	0.0000	0.0000	0.0000
GRGDP	0.0006	0.3827	0.0003
LNGDPCAP	-	-	-
GRGNP	-	-	-
INTR_10Y	-	-	-
INTR_10Y2	0.2820	0.3028	0.2114
INTR_3M	-	-	-

INTR_3M2	-	-	-
SLYC	-	-	-
SLYC2	-	-	-
GRM1	0.0070	0.0000	0.0017
GRM2	0.0265	0.0871	0.0130
HPI	-	-	-

Notes: Column (1) refers to the test of the equality of coefficients in the three sub-periods. Column (2) reports the test of the equality of coefficients between pre-crisis and crisis period. Column (3) shows the test of the equality of coefficients between the crisis period and the post-crisis period.

Table A.4 Accuracy ratios and the number of surviving variables in the CPH model and its penalized versions

Model		CPH model	Penalized CPH model		DH model	Penalized DH model		Logistic model	Penalized Logistic mode	
			AIC	BIC		AIC	BIC		AIC	BIC
AUC	In-sample	0.238	0.779	0.779	0.761	0.749	0.749	0.762	0.755	0.678
	Out-of-sample	0.214	0.790	0.795	0.465	0.560	0.560	0.486	0.471	0.577
AR	In-sample	-0.524	0.559	0.559	0.523	0.497	0.497	0.523	0.511	0.356
	Out-of-sample	-0.572	0.581	0.591	-0.071	0.119	0.119	-0.028	-0.057	0.154
BS	In-sample	0.834	0.122	0.122	0.098	0.103	0.103	0.098	0.100	0.111
	Out-of-sample	0.771	0.118	0.118	0.223	0.199	0.199	0.219	0.211	0.178
Surviving variables		64	36	23	64	43	4	64	46	8

Notes: CPH model represents the Cox proportional hazard model and DH model refers to the discrete hazard model. “AUC” refers to the area under receiver operating characteristic curve. “AR” stands for accuracy ratio. “BS” represents the brier score. “AIC” is the AIC-type tuning parameter selector. “BIC” is the BIC-type tuning parameter selector.

Table A.5 IPO decision by out-of-sample prediction decile

Decile	CPH model	Penalized CPH model		DH model	Penalized DH model		Logistic model	Penalized Logistic model	
		AIC	BIC		AIC	BIC		AIC	BIC
1	0	64.22%	64.22%	13.76%	17.43%	17.43%	15.60%	15.60%	20.18%
2	0	4.59%	5.00%	2.75%	11.01%	11.01%	2.75%	3.67%	9.17%
3	0	4.59%	4.59%	11.93%	10.09%	10.09%	10.09%	12.84%	11.01%
4	0	5.50%	5.09%	7.34%	10.09%	10.09%	11.01%	7.34%	10.09%
5	0	3.67%	3.67%	10.09%	8.26%	8.26%	9.17%	9.17%	7.34%
6-10	100	17.43%	17.43%	54.13%	43.12%	43.12%	51.37%	51.37%	42.2%
AUC	0.214	0.790	0.795	0.465	0.560	0.560	0.486	0.471	0.577

Notes: CPH model represents the Cox proportional hazard model. DH model refers to the discrete hazard model. “AUC” refers to the area under receiver operating characteristic curve. “AR” stands for accuracy ratio. “BS” represents the brier score. “AIC” is the AIC-type tuning parameter selector. “BIC” is the BIC-type tuning parameter selector.

Section B

Accuracy ratios

AUC is a non-parametric measure generated from the receiver operating characteristic curve, which is commonly employed to assess the ability of a model to discriminate between binary events. It is already applied in related studies to evaluate the predictive ability to identify a default event (see for example Duffie et al., 2007 and Tian et al., 2015). The receiver operating characteristic curve is the plot of the likelihood of verifying true-positive (in practice, a bank issues IPO and the model classifies it as an expected event) and false-positive (in practice, a bank issues IPO but the model classifies it as an expected non-event) for a whole range of probable threshold points of probability values. If AUC is equal to 1, it represents a perfect prediction. If AUC is equal to or less than 0.5, it means that the corresponding model had no predictability. If the value of AUC is above 0.8, the predictive ability may be considered to be accurate (Hosmer Jr et al. 2013). The accuracy ratio is defined as the double difference between the value of AUC and 0.5, which is a frequently applied measure for corporate bankruptcy model evaluation. Thus, a value of 1 for accuracy ratio illustrates a perfect forecast, while a value of 0 for this shows a random forecast. To confirm the conclusions from AUC and the accuracy ratio, the brier score is included, to measure how close the predicted probability of a bank issuing IPOs in order to go public is to a bank staying in the private market. It is equal to the average of the squared differences between the forecast probabilities and the actual outcomes (1 if a bank issues IPO and 0 if a bank does not issue it). The brier score can be expressed as $\frac{1}{N}\int_{t=1}^{N}(p_t - o_t)^2$, where p_t is the forecast probability of a bank issuing IPO and o_t is the corresponding actual event. The lower the brier score is for a series of predictions, the better the predictions are deemed to be.