T.C.
KADIR HAS UNIVERSITY
GRADUATE SCHOOL OF SOCIAL SCIENCES
FINANCE AND BANKING

An Empirical Study on Credit Early Warning Systems

By
Haluk Öngören

PhD Thesis

Thesis Supervisor
Professor Dr. Nurhan Davutyan

January, 2016

Approval of the thesis

AN EMPRICAL STUDY ON CREDIT EARLY WARNING SYSTEMS

Submitted by Haluk Öngören in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Finance and Banking, Kadir Has University by,

Professor Dr. Nurhan Davutyan Supervisor, Kadir Has University	
Professor Dr. Gürbüz Gökçen Marmara University	
Professor Dr. Osman Zaim Kadir Has University	
Associate Professor Dr. Eray Yücel Kadir Has University	
Assistant Professor Dr. Barış Altaylıgil Istanbul University	

Approval Date : 21st of January, 2016

“Me, Haluk Öngören, I declare that this PhD Thesis contains no material that has been accepted for the award of any other degree or diploma in any educational institution and, to the best of my knowledge and belief, it contains no material previously published or written by another person, except where due reference is made in the text of the thesis.”

Haluk Öngören

Acknowledgement

I would like to express my special appreciation and thanks to my supervisor Professor Dr. Nurhan Davutyan, you have been a tremendous mentor for me. It would have been almost impossible to finish without your guidance and deep understanding. Thank you!

I would also thank to my supervising committee members, Professor Dr. Osman Zaim, Associate Professor Dr. Eray Yucel for your very useful comments and suggestions. Thank you!

Thanks are also due to my defense committee members Professor Dr. Gürbüz Gökçen and Assistant Professor Dr. Barış Altaylıgil for their helpful encouragements.

A special thanks to my family. I really don't know how to express my deep gratitude to you. Words are not enough to explain, how grateful I am to my mother and father, you have raised me up. My wife, your love, support and patience have smoothed my way through out my long journey. You have stood by me along this road with your love and support. Thank you Zeliha, for sharing the life.

Thank you Fatih Emre, thank you Ali Talha my little boys since some of your playtime has been sacrificed for my research.

Thank you all!

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Technical Terms

AMEX	American Stock Exchange
BRSA	Banking Regulation and Supervision Agency
CB	Conventional Bank
CBRT	Central Bank of the Republic of Turkey
COBIT	Control Objectives for Business IT
Due Diligence	An investigation of a business or person prior to signing a contract
EAD	Exposure at Default
FDIC	Federal Deposit Insurance Company
ICH	Interbank Clearing Houses
IT	Information Technologies
KSE	Korean Stock Exchange
LGD	Loss Given Default
LLS	Log-log Survival
Mitigant	Devices such as collateral, pledges, insurance, or guarantees that are used to decrease the credit risk exposure
NPL	Non-Performing Loan

NYSE	New York Stock Exchange
Obligator	Borrower, obligor
Obligee	Creditor
PBs	Participation Banks
PD	Probability of Default
Profit Share Rates	This is the term used by participation banks to describe what they pay to their depositors.
SFHs	Special Finance Houses
SME	Small and Medium Enterprises
TDR	Term Deposit Rates

ABSTRACT

AN EMPIRICAL STUDY ON CREDIT EARLY WARNING SYSTEMS

Haluk Öngören
Graduate School of Social Sciences Finance and Banking
Thesis Supervisor : Professor Dr. Nurhan Davutyan
January, 2016

Due to its impact on profitability and its potential regulatory consequences, financial distress prediction is vitally important for banks. The first generation of prediction models were based on the dichotomous classification of survival versus failure states and utilized balance sheet figures, and income statements of bank customers to make predictions. However those models were not designed to accommodate the change in the financial situation of bank customers over time.

We define *default* broadly as the bank declaring a loan as non-performing or initiating the legal process to collect the claimed amounts from the borrower. In this study, we use Cox's PH – Proportional Hazard approach to predict the potential defaulters using an unbalanced panel data set from 2005 and 2012. We have 202,615 observations on 15,593 customers obtained from one of the most reputable participation banks.

To our knowledge it is the first application of the Cox PH model to predict financial distress of bank borrowers. It is also important to note that it is also the first such study where only core banking information namely accounting and lending records is used. We did not adopt the traditional approach and thus did not use customer financial statements in our study.

We create three different financial distress models and use *selectivity ratio* and *success rate for defaulters* terminology to analyze which model's predictive performance is better. We conclude that, 72.41% of actual defaulters in the first quarter of 2013 and 58.37% of actual defaulters in 2013 have already been predicted by our Model at the end of 2012.

Key Words : Financial distress, early warning systems, Cox proportional hazard model, credit risk.

ÖZET

KREDİ ERKEN UYARI SİSTEMLERİ ÜZERİNDE
AMPİRİK BİR ÇALIŞMA

Haluk Öngören
Sosyal Bilimler Enstitüsü, Finans ve Bankacılık Doktora Programı
Danışman : Prof.Dr. Nurhan Davutyan
Ocak, 2016

Mali sıkıntı yaşayan müşterilerin tespiti gerek banka kârı, gerekse de regülasyonlar açısından çok önemli bir konsepttir. Birinci jenerasyon tahminleme modelleri bilanço ve gelir tabloları gibi finansal tablolardan elde edilen dikotom (ikili) değişkenler üzerine kurgulanmış olup zaman boyutunda değişkenleri ihtiva etmemektedir.

Borçlunun iflası tabiri ile müşterinin kredisini geri ödeme kapasitesini kaybetmesi ve/veya iflas hali ile birlikte banka tarafından alacağın nakde dönüştürülmesi ile ilgili hukuki süreçlerin başlatıldığı durum ifade edilmektedir. Bu çalışmada, saygın bir Katılım Bankası'ndan 2005 – 2012 yılları arasını kapsayan 15,593 farklı müşteriye ait 202,615 gözlem datası kullanılarak Cox PH – Proportional Hazard yöntemi ile iflasa meyilli olan borçlular önceden tespit edilmeye çalışılmıştır.

Bildiğimiz kadarıyla, bu çalışma banka datası üzerinden Cox PH yöntemi kullanılarak iflasa meyilli müşteri tahminlemesinin yapıldığı ilk çalışmadır. Ana bankacılık sisteminden alınan müşteri hesap ve kredi kayıtları ile yapılan ilk çalışma olduğunu da belirtmek gerekir. Geleneksel yöntemlerden farklı olarak müşterinin finansal tabloları çalışmamızda kullanılmamıştır.

Çalışmamızda üç farklı tahmin modeli geliştirdik ve modellerimizin birbirine karşı tahmin performansını ölçümleyebilmek için de "*seçicilik rasyosu*" ile "*iflas tutturma oranı*" adını verdiğimiz iki değişken kullandık. 2012 yıl sonu datasını kullanarak Model 3 ile yaptığımız tahminlemede 2013 yılının ilk üç ayında iflas eden müşterilerin %72.41'inin, 2013 yılında iflas eden müşterilerin ise %58.37'sinin modelimiz tarafından önceden tahminlendiğini gördük.

Anahtar Kelimeler : Mali sıkıntı, erken uyarı sistemi, Cox proportional hazard modeli, kredi riski.

1. Participation Banking

1.1. Introduction

The financial crisis in 2008 has demonstrated the necessity for banks of investing more in credit monitoring. Information on borrower quality is a key resource for lenders. Information can be from borrower's financial statements as well as unstructured information from news, magazines or social media. There is also the behavioral information of the borrower. As the use of technology has entered banking after 1980s' the banks were able to reach millions of customers. A lot of information is accumulated in the core banking systems of banks containing invaluable information. The emergence of alternative distribution channels facilitating the execution of banking transactions makes life far easier for customers. As a result, financial depth has increased and the banking has entered everyone's life.

As banks are able to reach millions of customers it is a great challenge to monitor the borrower's default risk. It is clear that banks which only utilize borrower's financial statements are definitely missing the great trove of information accumulated in their core banking systems. In this study we investigate whether the accounting, credit line usage and other structured set of information in core banking and its satellite systems can be useful in financial distress prediction and conclude they certainly are useful. We believe this information is very valuable because it enables the bank to cheaply utilize huge amounts of data. Note that existing practice in many financial institutions require loading financial statements collected from thousands of customers. Second, our approach not only dispenses with the costs involved with such data collection and loading, it uses ***up to date*** customer information already in the bank's data environment. Third, as opposed to customer financial statements that can be inaccurate[1], we make use of the bank's own records that are free of biases. We believe this approach will not only serve banks in maximizing their profits but also help borrowers to better sense approaching financial distress so that appropriate precautions can be taken.

We used the Cox PH – Proportional Hazard method in our study. This method is commonly used in health sciences for exploring the relationship between the survival of a patient and several explanatory variables. Cox model provides an estimate of the treatment effect on survival after adjusting for other explanatory variables. It allows to estimate the hazard or risk of death for an individual. In our study we make an analogy. We replace borrowers with patients, medical institutions with banks, explanatory variables with information received from core banking systems and patient death risk with borrower's default risk.

This study proceeds as follows. In Chapter 1 we give the idea of Participation Banking mainly its history, some figures about the market size and growth rate and basic terminology. In Chapter 2 we describe the credit decision. We explore the credit components and provide information about the lifecycle of lending activities. In Chapter 3 we survey the literature related to financial distress studies and in Chapter 4 we give information about the framework and our data. In Chapter 5 we present our findings and Chapter 6 concludes.

---

[1] See the informal sector discussion in Section 2.4.

1.2. Brief History of Participation Banking in Turkey

The foundation of PBs formerly known as ‘‘Special Finance Houses (SFHs)’’ was first approved in 1983. The first two Special Finance Houses - SFHs, Albaraka Turk and Faisal Finans, started their interest - free financial operations in 1985. In principle, they took deposits on the basis of profit and loss sharing, their depositors participated in an investment pool whose returns would not necessarily be positive. If the bank’s activities namely loans resulted in losses, deposits would shrink, some part of this loss would be allocated as loss to their depositors, i.e., they would get negative profit. Based on the premise that the depositors of finance houses accepted downside risk, their deposits were not insured. They were to earn not interest but variable returns based on the profitability of the projects financed.

Several finance houses were established on the following years, the largest was Ihlas Finans which went to bankruptcy during Turkish 2001 financial crisis. In the absence of thousands of depositors and investors lost money resulting massive withdrawals from all the finance houses where all the depositors and management of SFHs were panicked and tried to survive. As a lesson learned, the law governing their operations, Banking Law No. 5411 was revised in November 2005 and united the governance rules of SFHs with all the other Conventional Banks – CBs in Turkey and their deposits would be insured up to 50.000 TL. SFHs also were then called as “*Participation Banks*“.

The insurance limit is 100.000 TL as of now and currently, there are five participation banks in Turkey. There is still no special separate regulation for PBs, they are regulated and supervised by the same law by Banking Regulation and Supervision Agency - BRSA as conventional banks.

1.3. Need for Participation Banks

The desire for developing interest free banking system caused the formation of PBs as a kind of profit and loss sharing system. The idea is based on collecting funds without an agreed profit rate[2], and giving as loan to individuals with a reasonable profit ratio. This profit or loss is generally distributed by %80 to the account owner and %20 to the bank.

There two types of accounts in PBs.

i. Current accounts where the bank does not pay any return to the depositor.
ii. Time deposit accounts where PB pays a return to the depositor based on the realized profitability of the projects financed by the bank after deducting bank's share.

Note that there is no prearranged interest rate promised to the depositor. Thus such interest expenses are not a fixed cost for the PB. This is the major difference between commercial and participation banking from the deposit side. In other words, this is the practical meaning of interest free banking. Clearly this is quite advantageous for the management of PBs. On the other hand, the interest charged to the borrower is set up with the loan contract just as in commercial banking.

[2] However, note that although no explicit return promise is made to the depositors, there is an implicit promise. The proof of its existence comes from the bankruptcy of Ihlas Finans in 2001 due to its inability to make such payments. This point is made by Çokgezen and Kuran (2015).

1.4. Statistical Figures about Participation Banking

As of December 2015 there are five active PBs in the Turkish market.

Albaraka Turk Participation Bank, whose establishment was completed in 1984, became operational at the beginning of 1985. Founded as a joint undertaking between the Albaraka Banking Group (ABG), the Islamic Development Bank (IDB) and a Turkish industrial group that has been serving the national economy for more than half a century, Albaraka Türk boasts a strong capital base. As of 30 June 2014, 66.10% of the Bank's shares were held by foreign shareholders and 10.48% by local shareholders while the remaining 23.42% were publicly traded.

Asya Katılım Bankası A.Ş. commenced its activities on October 24th, 1996, as the sixth special finance institution of Turkey. The company's name, which had been previously "Asya Finans Kurumu Anonim Şirketi", was changed to "Asya Katılım Bankası Anonim Şirketi" on December 20th, 2005. Bank Asya has a multi-partnered (195) structure based on wholly domestic capital.

Kuveyt Turk, which was established in 1989 in the status of Special Financial Institution, became the third institution to join the sector. 62% of the capital of Kuveyt Turk is owned by Kuwait Finance House, 9% by the Public Institution for Social Security of Kuveyt, 9% by the Islamic Development Bank, 18% by General Directorate for Foundations and 2% by other shareholders.

Türkiye Finans Participation Bank was established in 2005 with the merger of the Anadolu Finans and Family Finance institutions. 60% of the shares in Türkiye Finans were purchased by the most important bank in the Middle East and the largest bank in Saudi Arabia, The National Commercial Bank (NCB), on March 31, 2008.

Ziraat Participation Bank is the last player who entered the market in 2015. 100% of belongs to Ziraat Bank, the biggest public bank of Turkey. The governor party wants to increase the market share of the PBs, Ziraat PB is a part of that strategy.

In this part we tried to give an overview about the background and some statistical figures Participation Banking in Turkey.

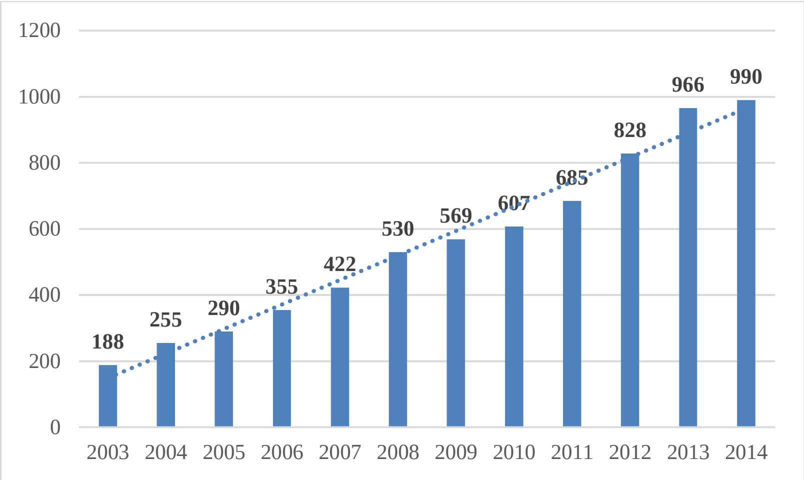


Figure 1 : Number of branches per years, Source: Participation Banks Association of Turkey

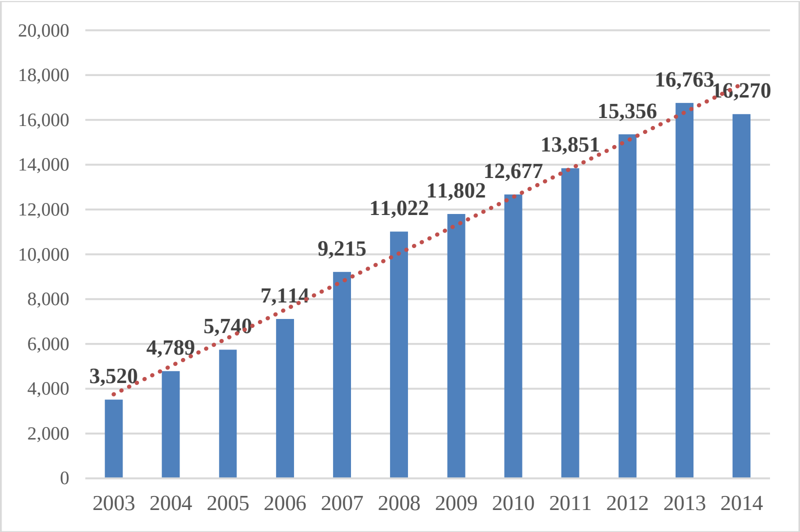


Figure 2 : Number of employees per years, Source: Participation Banks Association of Turkey

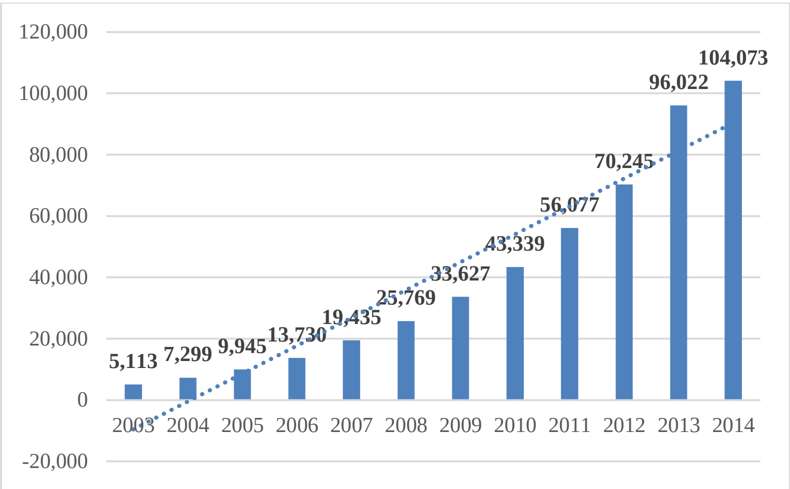


Figure 3 : Total assets per years (million TL), Source: Participation Banks Association of Turkey

Years	PBs (million TL)	Growth of PBs %	CBs (million TL)	Growth of CBs %	Market Share of PBs (%)
2000	2,266		106,549		2.08%
2001	2,365	4,37%	218,873	105.42%	1.07%
2002	3,962	67,53%	216,637	-1.02%	1.80%
2003	5,113	29,05%	254,863	17.65%	1.97%
2004	7,299	42,75%	313,751	23.11%	2.27%
2005	9,945	36,26%	406,915	29.69%	2.39%
2006	13,729	38,05%	498,587	22.53%	2.68%
2007	19,435	41,55%	580,607	16.45%	3.24%
2008	25,769	32,59%	731,640	26.01%	3.40%
2009	33,628	30,50%	833,968	13.99%	3.88%
2010	43,339	28,88%	1,006,672	20.71%	4.13%
2011	56,077	29,39%	1,217,711	20.96%	4.40%
2012	70,279	25,33%	1,370,614	12.56%	4.88%
2013	96,222	36,91%	1,750,000	27.68%	5.21%
2014	104,073	8,15%	2,000,000	14.29%	4.95%

Table 1 : Market share %, Source: Participation Banks Association of Turkey

It is clearly seen from Table 1 that the market share of PBs is constantly growing until the year 2013. We don't have figures before 2000 but it is commonly believed after the elections in 2002 since AK Party have begun to rule the country PBs gained credibility. Especially the regulation in 2005 which classified them as Participation Banks _i.e. the word "bank" was made part of their title_ thereby enabling governmental institutions to work with PBs was crucial. Note that prior to this change PBs were considered "special finance houses" which made it impossible for public agencies to deal with them since by law such agencies could only treat with "banks".

The peak year is 2013 and the PB sector's market share in the banking sector's total assets, which was 4.0% in 2009, reached to 5.2% by the end of 2014, with 25% average annual growth in its assets between 2009-2014. There happened to be a downsize in 2014 due to downsizing of Asya Katılım Bankası A.Ş. because of a disagreement with the governing party.

In the Turkey's Participation Banking Strategy study, with the new players in the participation banking system, the system's total assets' share in the total banking sector is expected to reach 15% in 2025[3].

Ongena and Yuncu (2011) have shown that PBs mainly deal with young, and transparent firms that are manufacturing and industry focused. Dolgun and Turhan (2014) argue that, PBs expand the scope for financial inclusion of people who stay away from conventional banking due to religious sensitivity. They also imply compared to commercial ones, participation banks are less likely to finance consumption loans. In this sense, they claim PBs play an important role in channeling idle capital to more productive sectors.

However, to date there is no comprehensive empirical work comparing the loan composition of participation versus commercial banks. Davutyan and Öztürkkal (2015) provide further clarification and some tentative evidence on these two issues.

[3] (http://www.tkbb.org.tr/documents/TKBB_Strateji_Belgesi_Ingilizce.pdf, 2015)

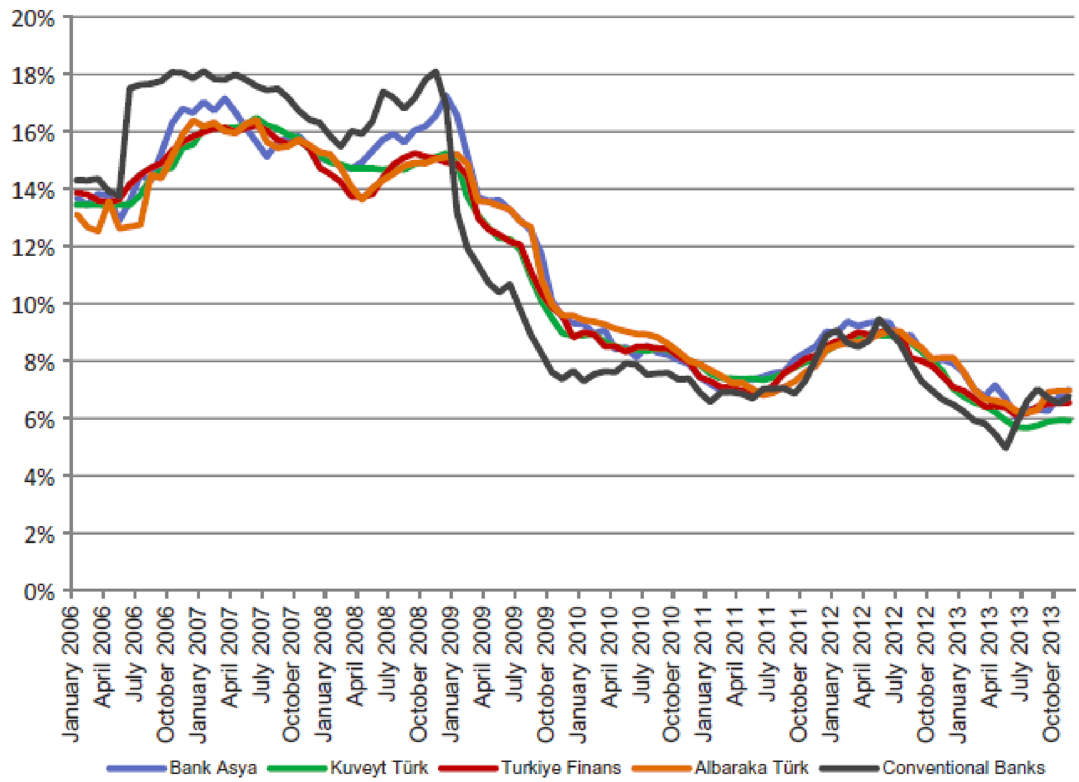


Figure 4 : Monthly change in annual TDR of participation banks and average annual TDR of conventional banks[4]

Like Islamic banks everywhere, Turkey's participation banks practice interest-free banking as defined previously. Officially, that is what differentiates them from commercial banks. Özsoy, Görmez and Mekik (2013) and Kaya (2013) have shown that this is one of the most important reasons why their customers choose PBs.

PBs are always criticized as to why TDRs and loan rates in CBs usually move together. Kuran (2004) argues that PBs give and take interest routinely, their depositors receive returns that are nearly identical to the rates paid by CBs. There is no statistically

[4] (Sarac & Zeren, 2015)

significant difference between the returns of the two groups of depositors. In lending, too, the participation banks impose charges that are practically indistinguishable from interest.

Saraç, and Zeren (2015) have carefully analyzed 2002 to 2013 Turkish data on deposit rates of PBs. Their empirical results show the TDRs or profit share rates are significantly cointegrated with TDRs paid by CBs. They also argue this very close correlation between the two rates are inevitable in the modern world where conventional finance dominates and competition tends to equate risk adjusted rates of return globally. Like Çevik & Charap (2011) and Chong & Liu (2009), there are also some other studies indicating the profit share rates of PBs closely track those of CBs.

Ahmad (1993) argues the apparent similarities between CBs and Islamic Banks are simply a phase in the transition away from conventional banking. Similarly, Mirakhor (2009) claims the Islamic financial system is only in its early stage of development and is operating coexistent with the conventional system in a hybrid form in which many of its supportive institutional elements either do not exist or are weak and incomplete. However, Khan (2010) shows Islamic banks simply replace conventional banking terminology with terms from classical Arabic and offers near identical services to its clients at a higher cost. Since as argued by Çokgezen and Kuran (2015) PBs differ only cosmetically from CBs, we will shift to banking terminology instead of PB from now on.

2. The Credit Decision

In writing this chapter, I have extensively used the excellent work on the credit decision by the "*The Bank Credit Analysis Handbook*" by Jonathan Golin and Philippe Delhaise.

The word *credit* derives from the ancient Latin *credere*, meaning "to entrust" or to "believe". Over the intervening centuries, the sense of the term remained close to the original; lender, or creditors, extend funds-or "credit"-based upon the belief that the borrower can be entrusted to repay the sum advanced, together with the interest, according to agreed terms. This conviction rests upon two fundamental principles; namely, the creditors confidence that;

i. The borrower is, and will be, willing to repay the funds advanced
ii. The borrower has, and will have, the capacity to repay those funds.

The first premise generally relies upon the creditor's information (or the borrower's reputation), while the second is typically based upon the creditor's understanding of the borrower's financial condition, or a similar analysis performed by a trusted party.

2.1. Components of Credit Risk

Credit risk evaluation can be considered as answering a series of questions in four areas.

i. The Obligator's Capacity and Willingness to Repay
    - What is the capacity of the obligator to service its financial obligations?
    - How likely will she/he/it be to fulfill that obligation through maturity?
    - What is the type of the obligator and usual credit risk characteristics associated with her/his/its business niche?

- What is the impact of the obligator's corporate structure, critical ownership, or other relationships and policy obligations upon its credit profile?

ii. The External Conditions

- How the country risk (sovereign risk) and operational conditions, including systemic risk, impinge upon the credit risk to which the obligee is exposed?
- What cyclical or secular changes are likely to affect the level of that risk? The obligation (product): What is its characteristics?

iii. The Attributes of Obligation from Which Credit Risk Arises

- What are the inherent risk characteristics of that obligation? Aside from general legal risk in the relevant jurisdiction, is the obligation subject to any legal risk specific to product?
- What is the tenor (maturity) of the product?
- Is the obligation secured; that is, are credit mitigants embedded in the product?
- What priority (e.g., senior, subordinated, unsecured) is assigned to the creditor (obligee)?
- How the specific covenants and terms benefit each party thereby increasing the credit risk to which the obligee is exposed? For example, are there any call provisions allowing the obligator to repay the obligation early; does the obligee have any right to convert the obligation to another form of security?
- What is the currency in which the obligation is denominated?

- Is there any associated contingent/derivative risk to which either party is subject?

iv. The Credit Risk Mitigants

- Are any credit risk mitigants - such as collateral – utilized in the existing obligation or contemplated transaction? If so, how do they impact credit risk?
- If there is secondary obligator, what is her/his/its credit risk?
- Has there been an evaluation of the strength of the credit risk mitigation?

2.2. Willingness to Pay

Willingness to pay is a subjective attribute that nobody can know for sure. It is related to the borrower's reputation and apparent character. But it is unknowable in advance. Therefore, evaluation is necessary from the perspective of the lender. Hence a *qualitative* evaluation that takes into account information collected from various sources, face-to-face meetings are a customary part of the process of *due diligence*.

Walter Bagehot, the nineteenth-century British economic commenter put it well:

> *"A banker who lives in the* district, *who has always lived there, whose whole mind is a history of the district and its changes, is easily able to lend money there. But a manager deputed by a central establishment does so with difficulty. The worst people will come to him and ask for loans. His ignorance is a mark for all the shrewd and crafty people thereabouts."*

So, in credit analysis willingness to pay should be taken account. This requires giving serious consideration to the borrower's past behavior. It is still up to the lender to decide the extent of importance to be attached to a borrower's character.

2.3. Indicators of Willingness

Willingness to pay is difficult to evaluate. Judgments and the criteria on which they are based, are subjective in nature.

- Character and reputation
- Credit record
- Creditors' legal rights and the legal systems can be considered as the indicators of willingness to pay.

Firsthand knowledge regarding a prospective borrower's character is a real test for credit decision. Where direct familiarity is lacking, the borrower's reputation provides an alternative basis for ascertaining the obligor's disposition to make good on a promise. However exclusive reliance on reputation can be perilous. A dependence upon second-hand information can easily descend into so-called *name lending*. Name lending can be defined as the practice of lending to customers based on their perceived status within the business community instead of on the basis of facts and sound conclusions derived from a rigorous analysis of prospective borrowers' actual capacity to service additional debt.

Nowadays, far more data is available, as technology has developed and credit reference agencies have been set up to provide this kind of service. A borrower's payment record can be an invaluable resource for the lender. Nevertheless, one should keep in mind that, although the past provides some reassurance of future willingness to pay, it cannot be extrapolated into the future with certainty in any individual case.

Legal and regulatory infrastructure and concomitant doubts concerning the fair and timely enforcement of creditors' rights also impact the willingness to pay. The stronger and more effectual the legal infrastructure is, the better able a creditor will be to enforce a judgment against a borrower. Prompt court decisions or the long arm of the state will tend to predispose the nonperforming debtor to fulfill its obligations.

So as legal systems have improved – together with the evolution of financial analytical techniques and data collection and distribution methods– the attribute of *willingness* to repay has been increasingly overshadowed in importance by the attribute of *capacity* to repay.

2.4. Evaluating the Capacity to Repay: Science or Art?

Compared to *willingness* to pay, evaluating the *capacity* to pay involves a more quantitative measurement. Applying financial analysis will give the clue whether the borrower will have the ability to fulfill outstanding obligations as they come due. Evaluating an entity's capability to pay derived from its most recent and past financial statements forms the core of credit analysis.

There are three serious limitations of financial analysis.

i. The historical nature of financial data.

ii. The difficulty of accurately projecting financial strength based upon such data.

iii. The gap between financial reports and financial reality.

The gap between financial reporting and reality is a well-known phenomenon in Turkey. In countries where the size of the informal sector is larger these problems becomes more prevalent. By analyzing the food expenditure data Davutyan (2008) showed that officially reported national income in 2005 should be multiplied by about 1.25, to obtain the true national income.

2.4.1. The Historical Character of Financial Data

Being invariably historical in scope and covering past fiscal reporting periods financial statements are never up to date. Because the past cannot be extrapolated into future with any certainty, except perhaps in cases of clear insolvency and illiquidity, estimating capacity remains just that: an estimate or a sophisticated guess. Accurate financial forecasts are notoriously problematic, and, no matter how refined, financial projections are vulnerable to errors, omissions and distortions. Small differences may lead to huge disparities in the range of values over time.

2.4.2. Financial Reporting may not be the Financial Reality

First, rules of reporting and financial accounting are shaped by people and institutions whose perspective and interest may differ. Influences emanating from that divergence are apt to aggravate these deficiencies. Second, there is the question of how various accounting items are treated. The difficulty in making rules to cover every tiny transaction may lead to inaccurate comparisons or further deception or fraud. Thirdly, the need for interpreting financial statements requires different vantage points, experience, and analytical skills. This may result in a range of somehow differing conclusions to be drawn from the same data. Considering everything, financial scrutiny remains at the core of an effective credit analysis in spite of its limitations and subjective elements. The associated techniques are essential and invaluable tools for drawing conclusions about a firm's creditworthiness, and the credit risk associated with its obligations. But on balance given the above mentioned reasons the seemingly objective evaluation of financial capacity retains a significant qualitative, and therefore subjective, component.

Thus it is crucial not to place too much faith in quantitative methods of financial analysis for assessing credit risk, nor to believe that quantitative data or conclusions drawn from such data necessarily represent the objective truth. No matter how sophisticated, when applied for the purpose of evaluating credit risk, these techniques remain imperfect tools that seek to predict an unknowable future.

2.5. A Quantitative Measurement of Credit Risk

Given such shortcomings, the softer more qualitative aspects of the analytical process should not be ignored. Notably, a thorough evaluation of management-including its competence, motivation, and incentives-as well as the plausibility and coherence of its strategy remains an important element of credit analysis for both nonfinancial and financial companies. Indeed, not only is credit analysis both qualitative and quantitative in nature, but nearly all of its ostensibly quantitative aspects also have a significant qualitative dimension.

Thus evaluating the willingness to pay and assessing management expertise and ability comprise subjective judgments. Although it is often overlooked, the same applies to the presentations and analysis of a firm's financial results. Credit analysis is as much art as it is mathematical inquiry. The best credit analysis is a synthesis of quantitative measures and qualitative judgments.

2.6. Credit Risk Analysis versus Credit Risk Modeling

Here, it is important to note there is a critical distinction between credit risk analysis and credit risk modeling. For example consider the concept of rating migration risk[5]. It is an important factor in modeling and evaluating portfolios of debt securities. However, it does not concern the credit analyst performing an evaluation of the kind upon which its rating is based. It is important to recognize this distinction and to emphasize the aim of the credit analyst is not to model credit risk, but instead to perform the evaluation that provides one of the requisite inputs to credit risk models. Naturally, it is also one of the indispensable inputs to the overall risk management of a banking organization.

2.7. A Quantitative Measurement of Credit Risk

So far, our inquiry into the meaning of credit has stayed within the bounds of tradition. *Credit risk has been defined as the likelihood that a borrower will perform a financial obligation according to its terms; or conversely, the probability that it will default on that commitment*. The chance that a borrower will default on its obligation to the lender generally equates to the probability that the lender will suffer a loss. *As so defined, credit risk and default risk are essentially equivalent*. While this has long been an acceptably functional definition of creditworthiness, developments in the financial services industry and changes in the sector's regulation over the past decade have compelled market participants to revisit the concept.

---

[5] The risk that a portfolio's credit quality will materially deteriorate over time without allowing a repricing of the constituent loans to compensate the creditor for the now higher default risk being undertaken.

2.7.1. Probability of Default

If we think deeper about the relationship between credit risk and default risk, it becomes clear that such probability of default (PD), while highly relevant to the question of what constitutes a "*good credit*" and what identifies a bad one, is not the creditor's sole, or in some cases even her central concern. Indeed, a default could occur, but should a borrower through its earnest efforts remedy matters promptly- thus making good on the late payment through the remittance of interest or penalty charges-and resume performance without further violation of the lending agreement, the lender would be made whole and suffer little harm. Certainly, nonpayment for a short period might cause the lender severely significant liquidity problems, in case it was relying upon payment to satisfy its own financial obligations, but otherwise the tangible harm would be negligible. Putting aside for a moment the impact of default on a lender's own liquidity, if mere default by a borrower alone is not what truly concerns a creditor, what then is the real cause of worry?

2.7.2. Loss Given Default

In addition to the possibility of default, the creditor is, or arguably should be, equally concerned with the severity of the consequences that a default would entail. It is perhaps easier to comprehend retrospectively. Was it a brief, albeit material default, like that described in the preceding paragraph that was immediately corrected such that the creditor received all the expected benefits of the transaction?

Or was it the type of default in which payment and no further revenue is ever obtained by the creditor, ending in a substantial loss as a result of the transaction? Obviously, all else being equal, it is the possibility of the latter that most worries the lender.

Both the probability of default and the severity of the resulting loss in case of default—each being conventionally expressed in percentage terms—are crucial in determining the tangible expected loss to the creditor. Of course there is also the creditor's understandable level of apprehension. The loss given default (LGD) summarizes the likely percentage impact, under default, on the creditor's exposure.

The third variable that needs considering is exposure at default (EAD). EAD may be stated either in percentage of the nominal amount of the loan (or the limit on a line of credit) or in absolute terms.

The three variables—PD, LGD, and EAD—when multiplied, give us expected loss for a given time horizon.

It is straightforward to see all three variables are quite easy to calculate after the fact. Examining its entire portfolio over a one-year period, a bank may determine that the PD, adjusted for the size of the exposure, was 5 percent, its historical LGD was 70 percent, and EAD was 80 percent of the potential exposure. Leaving out asset correlations within the loan portfolio and other complications, expected loss (EL) is simply the product of PD, LGD, and EAD.

EL and its constituents are, however, much more difficult to estimate in advance. Again past experience may provide some guidance.

All the foregoing factors are time dependent. The longer the *tenor* or duration of the loan, the greater the chance that a default will occur. EAD and LGD will also change with time, the former increasing as the loan is fully drawn, and decreasing as it is gradually repaid. Similarly, LGD can change over time, depending upon the specific terms of the loan. The nature of the change will depend upon the specific conditions and structure of the obligation.

2.7.3. Application of the Concept

To summarize, expected loss is fundamentally dependent upon four variables, with the period often taken to be one year for purposes of comparison and analysis. On a portfolio basis, a fifth variable, correlation between credit exposures within a credit portfolio, will also affect expected loss.

The PD/LGD/EAD concepts just described are very valuable as a way to understand and model credit risk.

As can be seen credit risk modeling framework is rich enough to encompass many different concepts. But our study is mainly involved with probability of default and time horizon. Thus in our empirical work we shall mainly utilize the latter two: probability of default and time horizon.

2.8. The Lending Process

Operating a successful lending business is more complicated than it might first appear. To make the activity profitable, funds must be sourced at a reasonable cost to lend to financially sound borrowers. Sourcing funds usually means attracting new depositors or attracting new deposits from existing depositors. Making sound loans necessitates identifying creditworthy loan applicants and projects. Both activities involve appropriate pricing on the one hand and cost-effective marketing on the other. With regard to lending, the prospective customer must be suitably approached or attracted to the institution. To this end, marketing is used to convey the proper image on the part of the bank and to educate the prospective customer concerning the benefits stemming from a relationship with the bank. Not surprisingly, pricing is another essential factor. Terms governing the loan agreement has to be acceptable both to the customer and to the bank. The major steps in the lending process are shown in Figure 5.

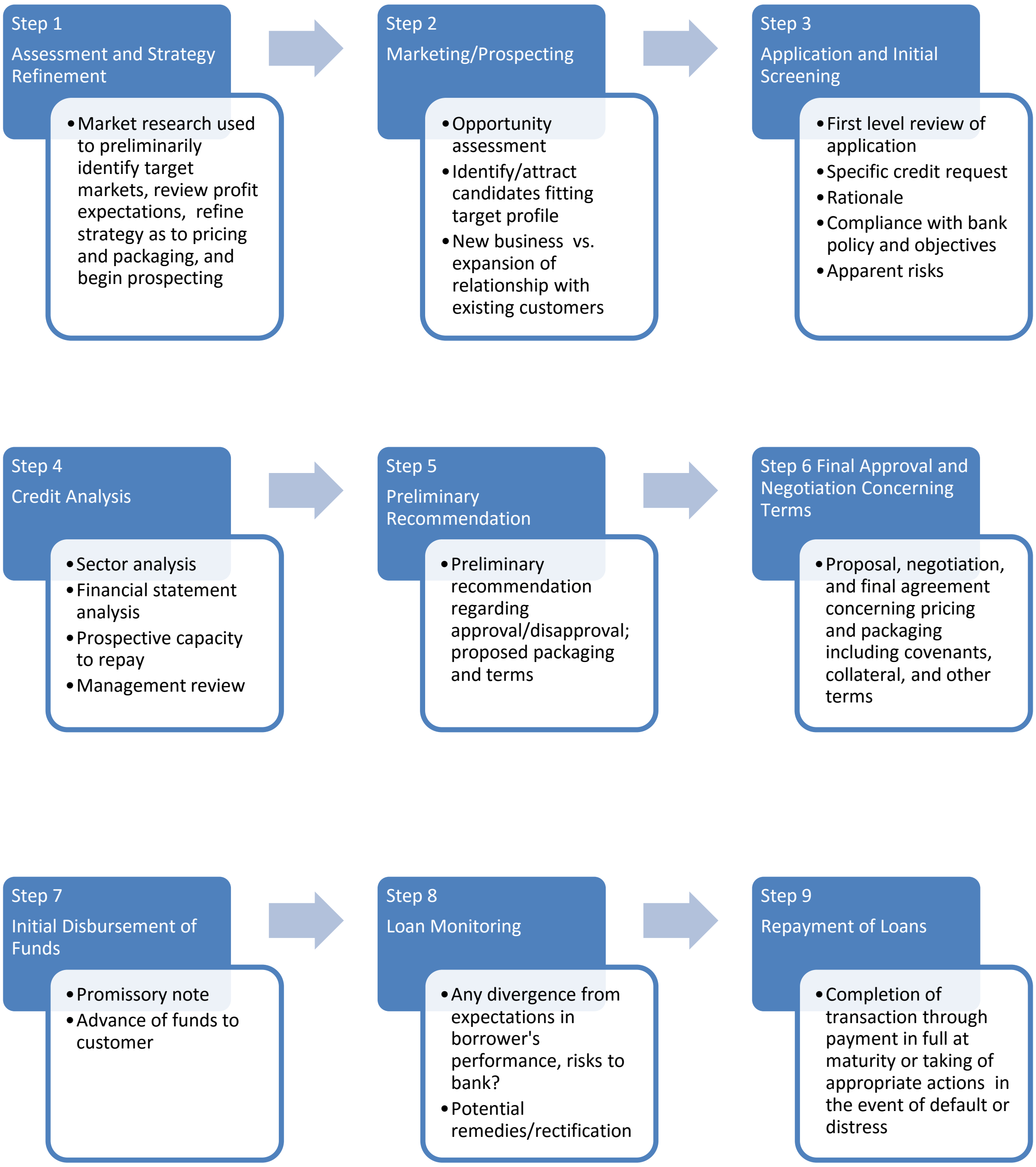


Figure 5 : The lending process[6]

[6] Source: "*The Bank Credit Analysis Handbook*" by Jonathan Golin and Philippe Delhaise p.105

The process begins with market research, the refinement of lending strategy and the formulation of tactics to attract the type of customers the bank seeks. Naturally, to this end (and also to appeal to depositors and facilitate inexpensive and stable funding), some sort of distribution network is needed. Traditionally, this has meant the development of a branch network to bring the bank closer to the customer, or as happened in the United States, the development of a highly localized banking system, which discouraged the creation of national branch networks by the largest institutions. Apart from collecting deposits, such a distribution network was also critical to developing a strong lending business. This follows from most customers' preference to deal with a locally accessible institution. The World Wide Web has brought with it e-banking. However not all types of lending activities are amenable to web distribution. Unsurprisingly banks that are purely web-based have found it more difficult than some expected to establish strong deposit networks.

Assuming a suitable distribution infrastructure has been established, the first operational step is to market the bank's lending and other financial services in order to attract desired loan applicants. With an application having been submitted, an initial review is performed to establish whether it broadly fits within the bank's guidelines as to whether the candidate is suitable, and whether risk levels look acceptable. This is followed by the process of packaging the loan to a particular applicant, and the initiating negotiations concerning pricing and loan terms including those governing collateral and covenants.

In the next phase of the process, credit analysis is performed, and a recommendation made. If affirmative, the proposal is considered by the appropriate credit committee. If approved, the agreement with the customer can be made. It is not uncommon for the credit committee to require some modification to the terms, particularly where large sums are at stake. If a final agreement is reached, this phase of the process concludes, the agreement is then formalized and funds advanced. The last phase of the process involves monitoring the customer and taking appropriate action in case of default or the emergence of new risks. Credit control staff assess and monitor collateral, while the bank's legal department keeps associated lending documentation on file and, in case of borrower distress, works out problematic loans. Finally, with the maturity of the loan and its repayment, the transaction concludes.

Jappelli and Pagano (2002) argue that information sharing among lenders attenuates adverse selection and moral hazard, and can therefore increase lending and reduce default rates. Using a new, purpose built data set on private credit bureaus and public credit registers, they find that bank lending is higher and credit risk is lower in countries where lenders share information, regardless of the private or public nature of the information sharing mechanism. They also find that public intervention is more likely where private arrangements have not arisen spontaneously and creditor rights are poorly protected.

In Turkey, the borrower credit line is reported to the central bank. The reporting is done on a quarterly basis. The information is accumulated in the Central Bank, and is then reported back to the bank(s), so that a bank is able to identify the total limit and exposure risk of any given borrower, including the number of banks that the borrower is working with. This information of great value to the bank, because it helps the bank to identify the overall risk of the borrower and assess the likely trajectory of pay back performance.

Apart from the Central Bank, private companies like FINDEKS[7] provide comparable information. They can be considered as credit reference agencies like Experian, Equifax and Call Credit PLC in England. As such they hold factual information on retail customers and this allows a lender to check individuals' names and address and past credit history, including any County Court Judgments or defaults recorded against the individual[8]. (Casu, Girardone, & Molyneux, 2006)

All standard procedures apply to PBs just in conventional banking. The process starts with the borrower applying to the bank for a credit limit. The borrower is evaluated according to his/her financial worth, financial performance, and according to a pre-defined set of religiously inspired principles. For instance, legally acceptable but morally objectionable activities, e.g. gambling, alcohol production and servicing, are not patronized.

If the PB decide to set work with the borrower it submits a terms and conditions letter which give the details about the guarantees and collaterals that the bank is requiring. In case the borrower agrees on the terms and conditions and he/she submits the pre asked guarantees and collaterals the PB opens the line of credit. The line of credit (i.e. limit of

---

[7] FINDEKS is a joint venture of 9 Turkish Banks founded in 1995 for information sharing of barrowers. As of today, they provide borrower information to more than 180 financial institutions including banks, consumer finance companies, insurance companies, factoring companies and leasing companies. https://www.findeks.com/kredi-kayit-burosu

[8] (Casu, Girardone, & Molyneux, 2006, p. 287)

credit) is defined as the maximum amount of money that a borrower can receive from the bank.

Once the credit limit is opened the borrower shall submit a valid reason to the bank justifying his need for credit. The reason can be either to buy a machine, raw materials, or acquiring any other physical goods from an external party (i.e. seller). It should be something tangible and real. The borrower cannot borrow for paying the salaries of his/her staff or for paying government imposed taxes to the tax office. The process starts with an application to the bank, the borrower should submit a pro forma invoice so that the process is compliant with the bank internal procedures. The bank verifies the pro forma invoice and after agreeing on the profit rate and maturity, an installments table is prepared with the given parameters by the internal rate of return calculation method. Once the installment table is generated, the capital is transferred (granted) to the seller (vendor) and the borrower is informed that the money has been paid; this process is referred to as a "project". Within 7 days the original invoice should be submitted to the bank. The amount of money that a borrower can request from the bank at any time is calculated by subtracting the sum of capital, profit amount, and tax from the approved credit limit. As the borrower pays out the installments, the amount of money that s/he can request for the next time is increased by the same amount. Hence, the overall risk is composed of different lending activities i.e. projects.

There is a minor difference in accounting between PBs and conventional banks, this difference arises from profit definition itself. It is important to note there is no special regulatory arrangement in terms of accounting principles belonging to PBs.

If we consider loan application, approval, collaterals and monitoring activities we can see that the processes and the methodologies followed are very similar with conventional banks. That's why in terms of our financial distress study or monitoring activities we believe there are no major differences between conventional and PBs hence our study applies both of them.

2.9. Loan Review

The conditions under which each loan is made change constantly. Such change affects the borrower's financial strength and her/his ability to repay. Fluctuations in the economy weaken some businesses and increase the credit needs of others, while individuals may lose their jobs or contract serious health problems, imperiling their ability to repay any outstanding loans. The loan department must be sensitive to such developments and periodically review all loans until they reach maturity.

While most lenders today use various loan review procedures, a few general principles are followed by nearly all lending institutions. These include:

i. Carrying out reviews of all types of loans on a periodic basis—for instance, routinely scrutinizing the largest loans outstanding every 30, 60, or 90 days, along with a random sample of smaller loans.

ii. Structuring the loan review process carefully to make sure the most important features of each loan are checked, including

   a. The record of borrower payments to ensure that the customer is not falling behind the planned repayment schedule.

   b. The quality and condition of any collateral pledged behind the loan.

c. The completeness of loan documentation to make sure the lender has access to any collateral pledged and possesses the full legal authority to take action against the borrower in the courts if necessary.
d. An evaluation of whether the borrower's financial condition and forecasts have changed, which may have impacted _upward or downward_ the borrower's need for credit.
e. An assessment of whether the loan conforms to the lender's loan policies and to the standards applied to its loan portfolio by examiners from the regulatory agencies.

iii. Reviewing the largest loans most frequently because default on these credit agreements could seriously affect the lender's own financial condition.

iv. Conducting more frequent reviews of troubled loans, with the frequency of review increasing as problems surrounding any particular loan increase.

v. Accelerating the loan review schedule if the economy slows down or if the industries in which the lending institution has made a substantial portion of its loans develop significant problems (e.g., the appearance of new competitors or shifts in technology that will demand new products and delivery methods).

Anecdotal evidence from corporate and consumer finance and from banking research indicate that offering a checking account along with a loan is important. By providing linked financial services, the bank can access information that is private, timely, quasi-costless, and reliable.

Mester, Nakamura and Renault (2007) show that transactions accounts, by providing timely and up to date data on borrowers' activities, help financial intermediaries monitor borrowers. This information is most readily available to commercial banks, which offer these accounts and lending together. They find that

i. Monthly changes in accounts receivable are reflected in transactions accounts;
ii. Borrowings in excess of collateral predict credit downgrades and loan write-downs; and
iii. The lender intensifies monitoring in response.

Norden and Weber (2010) argue that, in particular, the combined activity in a borrower's checking account and her/his credit line reveals significant information about her/his cash flow. That is, it provides the bank with information about the borrower's "debits" (*draws on the account that reflect cash outflows*) and "credits" (*receipts that reflect cash inflows*). Thus, these debits and credits may be the key determinant of a borrower's financial flexibility and debt repayment capacity (e.g., Sufi 2009). Unlike accounting numbers, payment data are less likely to be influenced by rules and policies.

However, account activity might be fragmented across different banks. This implies main banks would receive the greatest benefit from this source of information. Ongena and Yuncu (2011) state that PBs in Turkey mainly deal with multi-bank firms. Kaya (2013) have shown that 64% of PB customers work with CBs. Since PBs are usually not the main bank of their customers, this is an important disadvantage for those developing early warning systems based on such transaction accounts.

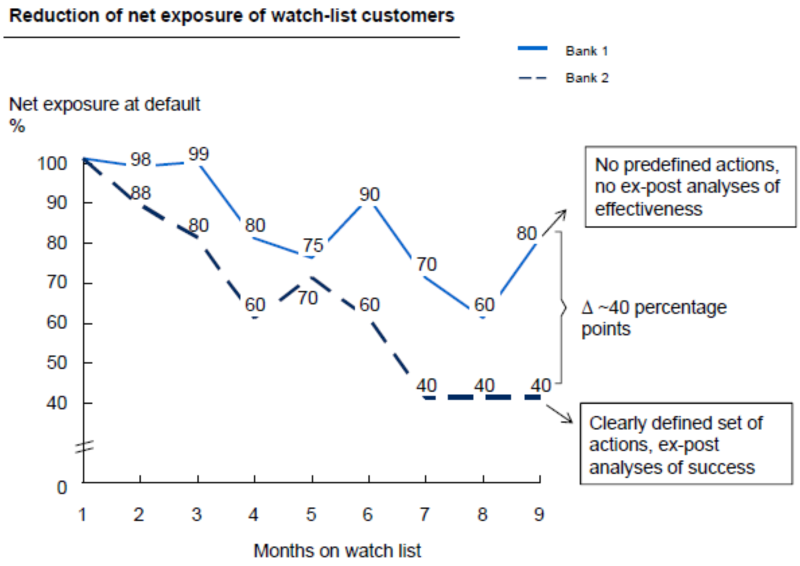


Figure 6 : Comparison of net exposure of two different banks

McKinsey (2012) argues banks with effective early warning systems identify risky customers six to nine months before they face serious problems, others may only take notice once a customer is past due or ratings have deteriorated substantially. Banks with good credit monitoring practices reduce unsecured exposures for customers on the watch list by about 60 percent within 9 months whereas average banks achieve only around 20 percent unsecured exposure reduction[9]. See Figure 6.

Regarding setting up early warning systems, there are mainly two approaches that banks can follow;

i. Develop the capability internally

---

[9] Available at http://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/risk/working%20papers/37_credit_monitoring_for_competitive_advantage.ashx

ii. Outsource this requirement

Some banks are trying to develop this capability internally. Since banks are generally huge organizations, the business departments assigned for this task face four main challenges;

i. The lack of business know-how and the relevant money & banking literature

ii. The lack of the necessary statistical knowledge.

iii. The need for a supporting IT infrastructure to allow the collection of information and integration with 3rd parties (IT vendors)

iv. Lack of coordination between different business units within banks – namely banks' internal organization prevents developing the capacity to predict customers' financial distress.

Hence, Turkish banks generally prefer to use less sophisticated models; crude rules of thumbs ("if then else" approach) are quite wide spread.

These services are also provided by international consulting companies. Experian, McKinsey, Fico and Oliver Wyman can be considered as examples who provides such services to their clients.

The software solutions provided by those vendors have five main disadvantages:

i. They are generally based on financial income statements, namely balance sheet and income statements. Collecting these documents from thousands of customers and input them into the system in a structured manner requires a hugely costly operational effort.

ii. The analysis based on such financial statements will not be accurate since window dressing is involved. The extent of informality in Turkey could give a conservative estimate of the amount of window dressing involved.

iii. The predictive model which is imported as a template needs at least three years of information collection before starting to generate results.

iv. They use logistic regression analysis which is less sophisticated than other available methods. We will be discussing these methods in Section 5.2.

v. Their high Total Cost of Ownership[10]

It is very important to transfer this capability from the vendor to the bank, otherwise the bank may not be able to operate or enhance the system without the help of the vendor.

Whatever approach that the bank follows, they should invest in;

i. The enhancement of the information flow between different business units to eliminate redundancy and any duplication of efforts, i.e. processes, roles & responsibilities

[10] Roughly speaking licensing costs around $1 million. When we include consultancy and implementation costs, the total would be around $3 million as initial setup cost. This can be considered as relatively high for a mid-sized Turkish bank. It is also worth noting that the bank would continue to pay around $200,000 yearly for maintenance.

ii. The necessary statistical knowledge to set up and interpret the results of any given econometric and mathematical analysis. These statisticians should also know and understand the banking environment.

iii. The enabling IT infrastructure.

Otherwise, the banks will not be able to judge the quality of the service as well as sustain the continuity of the required effort.

We have talked that the borrower is evaluated according to her/his financial worth, financial performance, and according to a pre-defined set of Islamic principles mentioned in Section 2.8.

A questionnaire consisting of three main parts namely behavioral, financial and nonfinancial questions are filled by credit analysts at least once each year for each customer. Depending on the difference of rating scores the bank may decide to;

i. Break off the credit relationship,

ii. Review terms and conditions and an increase or decrease in the credit line

iii. Continue to work on pre-agreed terms and conditions.

As the considerations listed above show, loan review is not a luxury but a necessity for a sound lending program. It not only helps management spot problem loans more quickly but also acts as a continuing check on whether loan officers are adhering to their institution's own loan policy. For this reason, and to maintain objectivity in the loan review process, most lenders separate their loan review personnel from the loan department itself. Loan reviews also aid senior management and the lender's board of directors in assessing the institution's overall exposure to risk and its possible need for more capital in the future.

2.10. Loan Workouts

In the natural course of business some loans will become problem loans despite all the safeguards that the bank builds. How often this occurs is related to borrower and project attributes as well as the general economic climate. Often such problems arise due to an excessive emphasis on the quantity rather than the quality of the loans booked[11].

The characteristic of each loan can be different but there are some common indicators of a weak or troubled loan.

The manual given to bank and thrift examiners by the FDIC discusses several telltale indicators of problem loans and poor lending policies:

Indicators of a Weak or Troubled Loan	**Indicators of Inadequate or Poor Lending Policies**
▪ Irregular or delinquent loan payments ▪ Frequent alterations in loan terms ▪ Poor loan renewal record (little reduction of principal when the loan is renewed)	▪ Poor selection of risks among borrowing customers ▪ Lending money contingent on possible future events (such as a merger)

[11] (Rose & Hudgins, 2008, p. 533)

▪ Unusually high loan rate (perhaps an attempt to compensate the lender for a high-risk loan) ▪ Unusual or unexpected buildup of the borrowing customer's accounts receivable and/or inventories ▪ Rising debt-to-net-worth (leverage) ratio ▪ Missing documentation (especially missing financial statements) ▪ Poor-quality collateral ▪ Reliance on reappraisals of assets to increase the borrowing customer's net worth ▪ Absence of cash flow statements or projections ▪ Customer reliance on nonrecurring sources of funds to meet loan payments (e.g., selling buildings or equipment)	▪ Lending money because a customer promises a large deposit ▪ Failure to specify a plan for loan liquidation ▪ High proportion of loans outside the lender's trade territory ▪ Incomplete credit files ▪ Substantial self-dealing credits (loans to insiders—employees, directors, or stockholders) ▪ Tendency to overreact to competition (making poor loans to keep customers from going to competing lending institutions) ▪ Lending money to support speculative purchases ▪ Lack of sensitivity to changing economic conditions

Table 2 : Warning signs of weak loans and poor lending policies[12]

[12] (Rose & Hudgins, 2008, p. 533)

Some potential leading indicators a borrower's likelihood of experiencing potential financial problems include:

i. Unusual or unexplained delays in receiving promised financial reports and payments or in communicating with bank personnel.
ii. For business loans, any sudden change in methods used by the borrowing firm to account for depreciation, make pension plan contributions, value inventories, account for taxes, or recognize income.
iii. For business loans, restructuring outstanding debt or eliminating dividends, or experiencing a change in the customer's credit rating.
iv. Adverse changes in the price of a borrowing customer's stock.
v. Losses in one or more years, especially as measured by returns on the borrower's assets (ROA), or equity capital (ROE), or earnings before interest and taxes (EBIT).
vi. Adverse changes in the borrower's capital structure (equity/debt ratio), liquidity (current ratio), or activity levels (e.g., the ratio of sales to inventory).
vii. Deviations of actual sales, cash flow, or income from those projected when the loan was requested.
viii. Unexpected or unexplained changes in customer deposit balances.

What should a lender do when a loan is in trouble? Experts in loan workouts—the process of recovering funds from a problem loan situation—suggest the following steps:

i. Lenders must always keep the goal of loan workouts firmly in mind: to maximize the chances for full recovery of funds.

ii. Rapid detection and reporting of any problems with a loan are essential; delay and procrastination often worsens a problem loan situation.

iii. The loan workout responsibility should be separate from the lending function to avoid possible conflicts of interest for the loan officer.

iv. Loan workout specialists should confer with the troubled customer quickly on possible options, especially for cutting expenses, increasing cash flow, and improving management control. Precede this meeting with a preliminary analysis of the problem and its possible causes, noting any special workout problems (including the presence of competing creditors). Develop a preliminary plan of action after ascertaining the lending institution's risk exposure and the sufficiency of loan documents, especially any claims against the customer's collateral other than that held by the lender.

v. Estimate what resources are available to collect the troubled loan, including the estimated liquidation values of assets and deposits.

vi. Loan workout personnel should conduct a tax and litigation search to see if the borrower has other unpaid obligations.

vii. For business borrowers, loan personnel must evaluate the quality, competence, and integrity of current management and visit the site to assess the borrower's property and operations.

viii. Loan workout professionals must consider all reasonable alternatives for cleaning up the troubled loan, including making a new, temporary agreement if loan problems appear to be short-term in nature or finding a way to help the customer strengthen cash flow (such as reducing expenses or entering new markets) or to infuse new capital into the business. Other possibilities include finding additional collateral; securing endorsements or guarantees; reorganizing, merging, or liquidating the firm; or filing a bankruptcy petition.

Of course, the preferred option nearly always is to seek a revised loan agreement that gives both the lending institution and its customer the chance to restore normal operations. Indeed, loan experts often argue that even when a loan agreement is in serious trouble, the customer may not be. This means that a properly structured loan agreement rarely runs into irreparable problems. However, an improperly structured loan agreement can contribute to a borrower's financial problems and be a cause of loan default.

According to "*Regulation on the Procedures and Principles for Determination of Qualifications of Loans and Other Receivables by Banks and Provisions to be Set Aside*" *(Published in Official Gazette Nr. 26333 dated November 1, 2006)* loans and other receivables are classified into five groups. Banks have to classify and monitor their loans and other receivables with respect to recovery capabilities.

Group One – Loans of a Standard Nature and Other Receivables, this group includes classification of loans and other receivables

- For which payments are made on terms, no repayment problems are expected in the future and which are totally recoverable / collectible and signs of weakness has been detected.

Group Two – Loans and Other Receivables under Close Monitoring, this group includes classification of loans and other receivables

- which do not presently face any problems in respect of principal or interest payments but which require close monitoring due to reasons such as observation of negative trends in borrowers' payment capability or cash flow positions or expectations for occurrence of such things or the fact that credit users face substantial financial risks or
- of which the repayment is highly likely but also the collection of capital and interest payments is delayed for more than thirty days as of the day of their payment dates for several reasons, however which do not carry the condition of delaying time to be classified among Group Three

Group Three – Loans and Other Receivables with Limited Recovery means: this group includes classification of loans and other receivables

- for which debtors have suffered deterioration in their creditworthiness and credits have suffered weakness consequently or
- for which recovery of principal and interest or both delays for more than ninety days from their terms or due dates provided that this is no more than one hundred eighty days or,
- for which it is believed that recovery by banks of principal or interest or both would delay for more than ninety days from their terms or due dates due to reasons such as problems encountered by debtors over operating capital financing or additional liquidity creation.

Group Four – Loans and Other Receivables with Suspicious Recovery: this group includes classification of loans and other receivables

- for which repayment or liquidation is not considered likely or
- for which the delay of recovery of principal or interest or both from respective terms or due dates exceeds one hundred eighty days provided that this delay is no longer than one year.

Group Five – Loans and Other Receivables Having the Nature of Loss: this group includes classification of loans and other receivables

- for which it is firmly believed that recovery is not possible or
- for which recovery of principal or interest or both delays for more than one year from respective terms or due dates

Group Three, Group Four and Group Five are considered as non-performing loans according to the same regulation. The legal process generally starts when at the point where loans and other receivable are turning into Group Three from Group Two and the borrower is considered as defaulter.

2.11. Collaterals and Guarantees

Collateral refers to assets that are either deposited with a bank, conditionally assigned to the bank pending full repayment of the funds borrowed, or more generally to assets with respect to which the bank has the right to obtain title and possession in full or partial satisfaction of the corresponding financial obligation. Thus, the bank who receives collateral and complies with the applicable legal requirements becomes a secured creditor, possessing specified legal rights to designated assets in case the borrower is unable to repay its obligation with cash or with other current assets. If the borrower defaults, the bank may be able to seize the collateral through foreclosure and sell it to satisfy outstanding obligations. Both secured and unsecured creditors may force the delinquent borrower into bankruptcy.

The secured creditor, however, benefits from the right to sell the collateral without necessarily initiating bankruptcy proceedings, and stands in a better position than unsecured creditors once such proceedings have started.

Credit risk mitigants are devices such as collateral, pledges, insurance, or guarantees that are used to decrease the credit risk exposure to which a bank or creditor would otherwise be subject. The purpose of credit risk mitigants is to ameliorate _ partially or totally_ a borrower's lack of intrinsic creditworthiness and thereby diminish the bank’s credit risk. For instance, where the borrower is comparatively new or lacks detailed financial statements, a bank may require a guarantee from a well-established enterprise which is rated by major external agencies. In the past, such mechanisms were frequently used to reduce or eliminate the need for the credit analysis of a prospective borrower by substituting conservatively valued collateral or the creditworthiness of an acceptable guarantor for the primary borrower.

In modern financial markets, collateral and guarantees, rather than being substitutes for inadequate stand-alone creditworthiness, may actually be a requisite and integral element of the contemplated transaction. Their essential function is unchanged, but instead of remedying a deficiency, they are used to increase creditworthiness to give the transaction certain predetermined credit characteristics. In these circumstances, rather than eliminating the need for credit analysis, consideration of credit risk mitigants supplements it. Real-life credit analysis consequently requires an integrated approach to the credit decision, and typically requires some degree of analysis of both the primary borrower and of the impact of any applicable credit risk mitigants.

Since the amount advanced is known, and collateral can normally be appraised with some degree of accuracy, often through reference to the market value of comparable goods or assets, the credit decision is considerably simplified. By anticipating the need to consider the issues of the borrower's willingness and capacity, the question, what is the likelihood that a borrower will perform its financial obligations in accordance with their terms, can be replaced with one more easily answered, namely: "*Will the collateral provided by the prospective borrower sufficient to secure repayment*?"

As Roger Hale, the author of an excellent introduction to credit analysis, puts it: "*If a pawnbroker lends money against a gold watch, he does not need credit analysis. He needs instead to know the value of the watch.*"

A guarantee is the promise by a third party to accept liability for the debts of another in the event that the primary obligor defaults, and is a kind of credit risk mitigant. Unlike collateral, the use of a guarantee does not eliminate the need for credit analysis, but simplifies it by making the guarantor instead of the borrower the object of scrutiny.

Typically, the guarantor will be an entity that either possesses greater creditworthiness than the primary obligor, or has a comparable level of creditworthiness but is easier to analyze. Often, there will be some relationship between the guarantor and the party on whose behalf the guarantee is provided. For example, a father may guarantee a finance company's loan to his son for the purchase of a car. Likewise, a parent company may guarantee a subsidiary's loan from a bank to fund the purchase of new premises.

Where a guarantee is provided, the questions posed with reference to the prospective borrower must be asked again in respect of the prospective guarantor: "*Will the prospective guarantor be both willing to repay the obligation and have the capacity to repay it*?" These questions are summarized in Table 3.

		Binary (Yes/No)	**Probability**
Willingness to pay	Primary Subject of Analysis (e.g., borrower)	Will the prospective borrower be willing to repay the funds?	What is the likelihood that a borrower will perform its financial obligations in accordance with their terms?
Capacity to pay		Will the prospective borrower be able to repay the funds?	
Collateral	Secondary Subject of Analysis (Credit risk mitigants)	Will the collateral provided by the prospective borrower or the guarantees given by a third party be sufficient to secure repayment?	What is the likelihood that the collateral provided by the prospective borrower or the guarantees given by a third party will be sufficient to secure repayment?

Guarantees		Will the prospective guarantor be willing to repay the obligation as well as have the capacity to repay it?	What is the likelihood that the prospective guarantor will be willing to repay the obligation as well as have the capacity to repay it?

Table 3 : Key credit questions

In view of the benefits of using collateral and guarantees to avoid the sometimes a long and costly task of performing an effective financial analysis, banks and other institutional lenders traditionally have placed primary emphasis on these credit risk mitigants, and other comparable mechanisms such as joint and several liabilities when allocating credit. For this reason, secured lending, which refers to the use of credit risk mitigants to secure a financial obligation as discussed, remains a favorite method of providing financing.

In countries where financial disclosure is poor or the requisite analytical skills are lacking, credit risk mitigants circumvent some of the difficulties involved in performing an effective credit evaluation. In developed markets, more sophisticated approaches to secured lending such as repo finance and securities lending have also grown increasingly popular. In these markets, however, the use of credit risk mitigants is often driven by the desire to facilitate investment transactions or to structure credit risks to meet the needs of the parties to the transaction rather than to avoid the process of credit analysis.

With the evolution of financial systems, credit analysis has become increasingly important and more refined. For the moment, though, our focus is upon credit evaluation in its more basic and customary form.

The complaint that banks do too little screening and tend to rely excessively on collateral may be particularly relevant for small business. In the United States, approximately 40% of the small business loans and almost 60% of their value are guaranteed and/or secured with personal assets: see Ang, Lyn & Tyler (1995) and Avery, Bostic & Samolyk (1998). In Turkey a common complaint of policy makers involves too little project screening and too much reliance on collateral by banks. Davutyan and Öztürkkal (2015) provide information on this issue.

Collateral requirements are larger for small businesses in developing countries and in backward regions of developed economies Harhoff and Körting (1998) argue that small firms from the former East Germany tend to pledge collateral more often than their counterparts in the former West Germany. Jappelli and Pagano (2002) argue the same is true for Southern Italy.

We will be discussing in our findings related to collateral requirements in Subsection 5.4.4.

According to the "Regulation on Procedures and Principles for Determination of Qualifications of Loans and other Receivables by Banks and Provisions to be Set Aside" dated November 1, 2006 published by BRSA, banks have to classify and monitor guarantees.

Collateral Type 1 covers:

a. Cash funds, deposits, participation funds and gold depot accounts provided that pledge or assignment contracts are arranged and funds provided from repo transactions made in return for bonds, bills and similar securities issued or guaranteed by the Turkish Treasury Under-secretariat, Central Bank, Directorate of Privatization Administration and Mass Housing Administration in respect of payment thereof and Type (B) investment fund participation certificates, receivables of member businesses arising out of credit cards and gold reserves maintained with banks,

b. Transactions carried out with the Turkish Treasury Under-secretariat, Central Bank, Directorate of Privatization Administration and Mass Housing Administration and transactions made in return for bonds, bills and similar securities issued or guaranteed by these entities in respect of payment thereof,
c. Securities issued by or on the basis of surety of the central governments and central banks of OECD countries and guarantees and sureties to be issued by this Bank,
d. Guarantees and sureties of banks operating in OECD countries,
e. Securities to be issued by European Central Bank or the surety of this Bank and the guarantees and sureties to be issued by this Bank,
f. Sureties, guarantee letters, bill guarantees, acceptances and endorsements that the banks operating in Turkey will give within the limits of their credit limits.

Collateral Type 2 covers:

a. Precious metals other than gold
b. Shares quoted with the Stock Exchange and Type (A) investment fund participation certificates
c. Private sector bonds and asset based securities excluding those issued by debtors
d. Credit derivatives agreement providing hedging against credit risks
e. Assignment or pledge of assessed receivables of natural persons and legal entities from the public entities

f. Securities that may easily convertible into currency and valuable papers representing merchandise and all kinds of merchandise and property at sums not exceeding their market values
g. Mortgages of real estate at the Land Registry and mortgages of real estate built on allocated land provided that their values according to expertise reports are adequate,
h. Export documents based on shipping bill or transport papers, or insured within the scope of exportation loan insurance policy
i. Bills of exchange stemming from actual trading relations, which are received from natural persons and legal entities

Collateral Type 3 covers:

a. Pledges on commercial operations
b. Other export papers
c. Pledge on vehicles, pledge on lines of commercial vehicles and license plates for commercial vehicles
d. Pledges on airplanes or ships
e. Sureties by natural persons and legal entities enjoying creditability higher than that of debtors
f. Other promissory notes received from natural persons and legal entities

Collateral Type 4 covers the types of guarantees not covered by the first three groups

BRSA have classified the collaterals according to their ability to convert cash in which the customer may able to pay back the loans. We will be analyzing relationship between the collateral type and the probability of default in the upcoming sections.

2.12. The Role of Information Technologies and Enterprise Governance

The US nominal GDP per capita was about $406 in 1913[13]. You could buy a REO car for $1.095 if you had enough money. In 2008 having reached $46,892[14] the GDP per capita was almost 113 times bigger than it was in 1913. By way of contrast in 2008, the price of an average car being about $28,350, the increase was only 26 fold. These numbers suggest the real increase in purchasing power was at least four involved. In other words, an average American had to work for 2.71 years to buy a REO car in 1913 but only 0.61 years in 2008. Considering the driving and security superiorities of new cars, the improvement involved is even more drastic than the 2.71/0.61=4.44 ratio suggests. Brynjolfsson and Saunders (2010) document the productivity enhancements which make such spectacular improvements possible. One chief engine of productivity growth especially after 1980 is the so called "computer revolution" which made information technology (IT) a regular input to the production process. Oliner and Sichel (2002) argue the acceleration in labor productivity after 1995 was driven largely by the greater use of IT capital goods and by the more rapid efficiency gains in producing of IT goods.

Dedrick, Gurbaxani and Kraemer (2003) have shown that IT investments have significant positive impact on firm productivity. Another key finding is that although returns to IT investments are positive on average, there is a wide range of differing performances across companies, with some doing much better than others. Some of these differences

---

[13] Johnston and Williamson (2015) Retrieved from http://www.measuringworth.org/usgdp/

[14] (National Automobile Dealers Association, Monthly Sales Trends, 2009)

could be explained by idiosyncratic firm differences that result in different opportunities to employ IT productively. In addition, they also show, there is strong evidence that investments in organizational capital through management practices such as decentralized decision making, job training, and business process restructuring have a major impact on returns to IT investments. The value of IT investments needs to be seen in relationship to investments in such organizational capital, as the two are complementary. IT is not simply a tool for automating existing processes, but is more importantly an enabler of organizational changes that can lead to productivity gains.

We can conclude that, information is a key resource for all enterprises, and from the time that information is created to the moment that it is used, technology plays a significant role.

Successful enterprises have recognized that IT should be embraced like any other significant part of doing business. Boards and management—both in the business and IT functions—must collaborate and work together, so that IT is included.

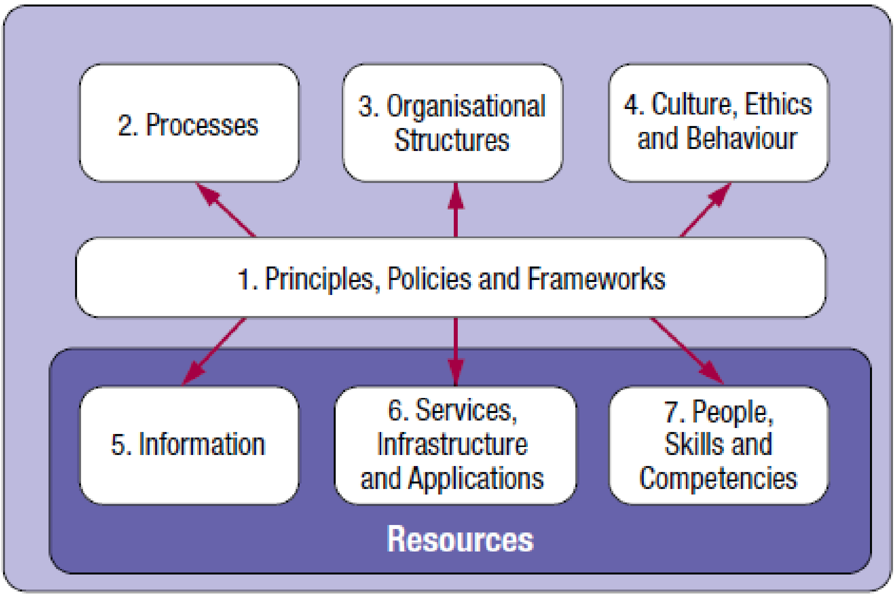


Figure 7 : COBIT 5 Enterprise enablers

COBIT5 – A framework for the governance and management of enterprise IT – suggests that enterprise must always consider an interconnected set of enablers. That is, each enabler:

i. Needs the input of other enablers to be fully effective, e.g., processes need information, organizational structures need skills and behavior.

ii. Delivers output to the benefit of other enablers, e.g., processes deliver information, skills and behavior make processes efficient.

In Turkey, this deficiency namely lack of an integrated approach within the business organization has been the rule especially after purchasing consultancy services or implementing new software. The risk manager of a one of the best known PB has once stated that, "*We use statistical techniques to calculate the risk scores of borrowers and report to BRSA. It is the business unit's own responsibility to decide whether to use or not to use the information that we have created within their business processes*". This statement implies there may be duplication of efforts within different business units and some information may not be consumed within the organization. Considering the huge organization involved _ e.g. thousands of staff and tens of thousands of customers, for an efficient i.e. waste free financial distress monitoring all above mentioned enablers should be taken into consideration in a bank.

3. Review of Related Literature

Beaver (1966) performed the first quantitative bankruptcy prediction analysis using a dichotomous classification test. He was followed by Altman (1968) who applied a z-score model based on discriminant analysis. Then came Ohlson (1980) and Zmijewski (1984) who adopted the logit and probit approaches respectively.

It will be noted that the above models are all based on a dichotomous classification between success and failure and thus are static. They have to split a collection of firms as those surviving vs those who went bankrupt without distinguishing between the differing calendar times of bankruptcy. But ignoring the time dimension involves throwing out valuable information because the failure process develops gradually over a long period of time. Therefore considering the time dimension of the firm's condition is a very important element in the bankruptcy modeling.

It follows that a statistical modeling approach with the ability to incorporate time will have serious benefits over the traditional static dichotomous approach. We can identify three advantages. First one can use time varying covariates at the firm level such as state of sales, staff, inventory, accounts receivable and payable, short and long term debt at various points in time. It should be stressed that the dichotomous approach is restricted to utilizing either time invariant characteristics such as region, sector, and owner's personal attributes or pick an *arbitrary* point in time and use economic and financial magnitudes pertaining to that moment. Second macro variables such as the state of the economy or the level of interest rates and the like are time varying. Thus the static dichotomous approach can only incorporate them at a particular point in time. This means they can only be used in cross country studies and are worthless for single country data.

Third, the presence of censored observations is important feature of real time data. Censoring occurs if the survival time of firms is longer than study period or the firms are excluded from observations for other reasons rather than financial distress.

Survival analysis is a statistical methodology designed to handle the time dependence of various real life processes. It has been applied in medicine, engineering as well as in social sciences like sociology.

Lane, Looney, and Wansley (1986) appears to be the first research paper applying survival analysis to default forecasting. Whalen (1991) examined bank failures and evaluated the usefulness of the proportional hazards model[15], a methodology that is going to be discussed in Subsection 5.2.1, as an early warning tool. The sample consisted of all U.S. banks that failed between Jan 1987 and Oct 1990, and randomly selected 1500 non-failed banks. They examined a particular type of early warning model called a Cox proportional hazards model, which basically produces estimates of the probability that a bank with a given set of characteristics will survive longer than some specified length of time into the future. Using a relatively small set of publicly available explanatory variables, the model identifies both failed and healthy banks with a high degree of accuracy. Furthermore, a large proportion of banks that subsequently failed are flagged as potential failures in periods prior to their actual demise. The classification accuracy of the model over time is impressive, since the coefficients are based on 1986 data and are

[15] To be discussed in Section 5.2

not updated over time. They concluded that the results demonstrate that reasonably accurate early warning models can be built and maintained at relatively low cost.

Chen and Lee (1993) focused on the failure of oil and gas industry to use Cox's proportional hazard model to predict financial distress. They applied survival analysis to study a class of financial distress when a financial analyst can identify an event that sets off the dynamic process of business adversity and would like to find out how long a firm can endure the adversity. They used the case of the oil and gas industry during the turmoil of the early 1980s and applied survival analysis to study how long a firm can endure this drastic oil price decline before facing financial distress. Their results showed that the liquidity ratio, leverage ratio, operating cash flows, success in exploration, age, and size are significant factors affecting corporate endurance.

Abdel-Khalik (1993) examined how well survival analysis predicts corporate financial distress in the oil and gas industry. The author feels that Chen and Lee do not give sufficient credit to previous research by W.R. Lane, S.W. Looney, and J.W. Wansley. The discussion dealt with the issues about experimental design and proportional hazard assumption, and stated that assuming the distribution rather than using the proportional hazard assumption was not justified. He also expresses reservations about the hazard-proportionality assumption and the information gained beyond an application of ordinary least squares.

Whalen (1991) and Chen and Lee (1993) treated the explanatory variables as time-invariant by fixing the values of the covariates at a given point in time. The major strength of proportional hazard model is that the model could employ time-varying covariates, but they didn't make full use of the benefit of survival analysis.

Wheelock and Wilson (1995) studied bank failures in Kansas from 1910 to 1928. Time-to-failure was explicitly modeled by using a proportional hazards framework. The results indicated that deposit insurance system membership increased the probability of failure, and technical inefficient banks were more likely to fail than technically efficient banks.

Helwege (1996) utilized a time-varying proportional hazard model to estimate the effect of asset allocation and funding choices on propensity of Saving & Loans to fail. Using data from financial reports filed with the Office of Thrift Supervision (OTS), time-varying proportional hazard functions are estimated to determine the extent to which failure was accelerated from increased usage of investment strategies since deemed riskier by Congress. Of particular interest is whether a greater concentration in whole residential mortgages and mortgage-backed securities (MBSs) greatly reduced the likelihood of failure over the 1980s. He argued that mortgage loans were significantly safer than other assets.

Shumway (2001) argued that hazard models are more appropriate than single-period models for forecasting bankruptcy. The model was developed using both accounting ratios and market-driven data for over 2000 companies from NYSE and AMEX over 31 years. Shumway concluded that the hazard model is theoretically preferable to static models in out-of-sample tests because it corrects for period at risk and allows for time-varying covariates. He found that half of the accounting ratios previously are poor predictors and several previously neglected market-driven variables are strongly related to bankruptcy probability.

Li-Sheng Chen (2005) employed the Cox model with time-varying variables to find the effect of biochemical covariates on death from liver cancer. They implemented a SAS

Macro program for time-dependent Cox regression predictive model for empirical survival data with time-varying covariates.

Kauffman and Wang (2008) used survival analysis to examine the factors of the internet business failures. They combined industry specific, business specific and macroeconomic variables as explanatory variables. The usefulness of the survival techniques showing that macroeconomic indicators are important elements was proved in this study.

Nam, Kim, Park and Lee (2008) investigated how the hazard rates of 367 listed companies in KSE are affected by changes in the macroeconomic environment and by time varying covariate vectors that show unique financial characteristics of each company. They also investigated out-of-sample forecasting performances of the suggested model and demonstrate improvements produced by allowing temporal and macroeconomic dependencies and found that the results of the out-of-sample forecasting showed that dynamic models with time-varying covariates are more accurate than static model. Among the dynamic duration models, a model with a macro-dependent baseline hazard rate showed superior performance. In fact, a direct specification of baseline hazard rate seems to be more important than any other factors, especially under the situation where the macroeconomic environment is changing abruptly and all the firms tend to be affected by the changes.

Oh, Nam, Kim & Lee (2013), Nam, Kim, Park & Lee (2008) assessed the violation of proportionality assumption in the firm failure prediction model built around the Cox's proportional hazard model and proposed non-proportional hazard model. They also examined the effect of macroeconomic variables to suggested non-proportional hazard model. They performed an investigation using the Korean stock market since the market, which has experienced two well-known structural changes caused by the Asian financial crisis and 2008 Global financial crisis, is well suited for analyzing the impact of the proportionality assumption on the appropriateness and predictability of the Cox's proportional hazard model.

All these studies are utilize publicly available data, and the financial distress studies are mainly related to bonds issued by firms quoted in stock markets. The major difference with our study is that, in our case we are trying to predict the financial distress of a bank's borrowers. Besides income statement and balance sheet, the bank has near real time information about the borrower's financial position. These transactions give a bank important information about the customer's cash flow and financial health. Obviously such information can be very valuable for predicting financial distress.

4. Institutional Framework and Data

4.1. Institutional Framework

Each customer has a unique customer identification number. The uniqueness of this number is guaranteed by controlling the tax number of the firm in question. This control is achieved through the core banking system, when the customer details are entered into the system.

Customers can have two types of accounts either time-deposit account or a current account. Each account is assigned to a branch i.e. branch identity (number), the customer identity and the account number are unique within the bank.

There are two major types of borrowing; it can be either cash or non-cash. However, there is only one type of credit limit. Non-cash borrowing includes all sorts of guarantees[16] which create a contingent liability for the bank. The sum of cash and non-cash borrowing should not exceed the pre-approved credit limit. In some cases, the limit is decreased or even set to zero due to some negative external information, or due to poor credit pay-back performance. In such exceptional cases the sum of the borrowed amounts (utilized credit limit) may exceed the limit itself.

---

[16] These guarantees were discussed in Section 2.11. In this case, the bank is guaranteeing that the seller will receive payment. In the event that the buyer –who is the customer of the bank- is unable to make the payment, the bank will cover the outstanding amount.

Each project has a unique ID and a profit rate, the redemption schedule based on the pre-agreed profit[17] rate at the beginning of each project.

The data we use comes from one of the most respected participation banks in Turkey. It involves analyzing credit line utilization, associated account movements, check clearing activities, demographic information, types and amounts of collaterals, import and export behavior of borrowing firms.

4.2. The data

Given this institutional framework, it is important to understand how different pieces of data are collected through the core banking system. Our study is mainly based on an unbalanced panel of 202,615 observations on 15,593 customers and 1,307 defaulters between the years 2005 and 2012. The length varies according to customer. The longest panel lasts 32 quarters, the shortest is 1 quarter. The panel on each firm ends with either an event occurrence; specifically a default or right censoring[18].

Data set that includes quarterly observations on limit usage, a review of account activities, and Central Bank data.

The panel data covers basically cover information related to;

---

[17] This is charged to the loan borrower and is partly distributed to depositors as we have already mentioned in Section 1.3.

[18] Such firms either stopped doing business with our bank for reasons other than failure between 2005–2012 or were in business at the end of 2012.

i. Firm characteristics (region, sector, date of foundation, amount of sales, number of staff, import/export dummies)

ii. Loan information (amount of credit line namely limit[19] and risk, limit utilization ratio, cash risk, collaterals)

iii. Accounting activities (quarterly average balance, quarterly minimum balance, #of debits, #of credits, etc.)

iv. Banking activities (amount of checks submitted to central bank, bounced checks, etc.)

While selecting firm data we have excluded individuals, financial institutions, charity foundations, all kinds of associations, unions, cooperative societies, and state or city owned entities, all official institutions and companies. Hence we can argue that each customer we have selected is either SME – Small or Medium Enterprise[20] or a corporate customer.

We start our study with the quarterly limit utilization data set extracted from the core banking system.

---

[19] Limit refers to the maximum amount of money that a borrower can receive from the bank as we have discussed in Section 2.8.

[20] As of 2012, SMEs are defined as enterprises either the number of staff is smaller than 50 or yearly amount of sales does not exceed 40 million Turkish Lira.

Second, we have included the limit risk usage details. The limit risk usage includes non-cash risk, cash risk, and credit limit, as well as customer identification number, in addition to the year/quarter data. The sum of non-cash and cash risk of exposure is labeled the total risk, the total risk divided by the limit is called the limit utilization ratio. The limit is set to be equal to total risk in case the total risk exceeds the limit due to reasons explained previously in Section 4.1. The panel data contains all borrowers whose total risk or limit is greater than zero in at least one of the given quarters, within the sample period. Thus, the data are restricted to borrowers who only received funding after 2005.

Check is a means of payment commonly used by SME's and corporate firms hence we have also included data from the central bank. Total size of checks sent and amount of returned/bounced checks sent for clearance to Interbank Clearing Houses - ICH of a given customer are also included in the data set.

Third, to take account of regional influences we use 7 regional dummies for Turkey's geographic regions:

Fourth, as a control for the general macroeconomic climate we use the overall percentage of the monetary value of bounced checks. This information is compiled by the Turkish Central Bank as part of its credit registry services. Note that

We operationalize the screening efforts of the bank by looking at;

i. Check Length,
ii. Check_NumberOfTransactions and
iii. Check_AmountOfTransactions.

For each firm, Check Length is the quarterly time difference between the date of first transaction in its account and the date of granting credit; Check_NumberOfTransactions is the number of all credit and debit transactions prior to granting of credit line; Check_AmountOfTransactions is the monetary sum of all debits and credits prior to opening a credit line. In every case a *larger* magnitude reflects a *greater* amount of information accumulation by the bank *before* making the credit decision.

We have also used data from the Central Bank's credit registry[21] pertaining to each customer's credit usage from other banks in Turkey. These are Total Limit_iob, Total Risk_iob, Limit Utilization Ratio_iob.

We distinguish between borrowers of various size. We use sales as a measure of borrower size. Specifically, we create three size ranges and create a classificatory dummy for each firm as DLarge, DMedium and DSmall on the basis of their sales.

Fourth, we quantified account(s) activities. We have calculated quarterly average balance using the end of day balance. Similarly, we calculated quarterly standard deviation of these daily balances, as well as number of credits and debits over the quarter. Since the customer may have different types of currency we have converted them all to Turkish Lira by using currency exchange rate announced by Turkish Central Bank on that day.

---

[21] As we discussed in Section 2.8 loan information related to every borrower is sent to the Central Bank by each bank. In return centrally consolidated information regarding each borrowing customer is made available to the relevant bank.

Table 4 describes the structure of our final data set.

Time Invariant Data based on Customer	Time Variant Data based on Year and Quarter
1. Customer based demographic data a. Sector [sector02 to sector18] b. Region [dreg2 to dreg7] c. Import & Export dummies [import, export] d. AgeAtLoan [age_at_loan] 2. If defaulted a. Default flag, b. Year of default, c. Quarter of default 3. Screening data a. Check length [check_length]: Quarterly time difference between the date of first transaction and date of opening credit line 4. Customer Type dummies a. dSales_l is 1 if 'Large' b. dSales_m is 1 if 'Medium' c. dSales_s is 1 if 'Small' sector01:unknown, sector02:fishing, sector03:social services, sector04:Education, sector05:electricity, gas and water resources sector06:real estate trading, sector07:financial intermediation, sector08:manufacturing, sector09:construction, sector10: sole proprietorship, sector11:mining and quarrying, sector12:hotels and restaurants, sector13:wholesale and retail trade, sector14:defense, public	1. Customer a. Number of staff [numberofstaffquarterly] 2. Assets a. Quarterly average balance [quarterlyaveragebalance] b. Number of debits [numberofdebits] c. Number of credits [numberofcredits] d. Returned Checks Percentage [rtrnedchecksprctn]: is the ratio of sum of bounced checks to total checks sent to ICH 3. Loans a. Total risk [totalrisk] b. Total limit [totallimit] c. Limit Utilization Ratio [limitutilizationratio]: is the ratio of risk to the total limit d. CashRiskRatio [cashriskratio]: is the ratio of cash risk to the total risk e. Total limit in other banks [totallimit_iob]: Sum of all limits in all banks excluding bank x; where bank x is the focus of analysis in this thesis. f. Total risk in other banks [totalrisk_iob]: Sum of all risks limits in all banks excluding bank x; where bank x is the focus of analysis in this thesis. g. Limit Utilization Ratio in other banks [limitutilizationratio_iob]: is the ratio of total risks to the

administration and social security, sector15:agriculture, hunting and forestry, sector16:transportation, warehousing, communication, sector17:small enterprises, sector18:health and social services dreg1:Akdeniz, dreg2:Doğu Anadolu, dreg3:Ege, dreg4:Güneydoğu Anadolu, dreg5:Iç Anadolu, dreg6:Karadeniz, dreg7:Marmara	total limits in all banks excluding bank x; where bank x is the focus of analysis in this thesis. 4. Collaterals a. w_collateral_1: Ratio of CollateralType1 to the sum of CollateralType1 to CollateralType4 b. w_collateral_2: Ratio of CollateralType2 to the sum of CollateralType1 to CollateralType4 c. w_collateral_3: Ratio of CollateralType3 to the sum of CollateralType1 to CollateralType4 5. Macroeconomic variable a. Central Bank returned cheques percentage

Table 4 : List of explanatory variables used in our models

	Small Firms		Medium Firms		Large Firms	
	Mean	**SDev**	**Mean**	**SDev**	**Mean**	**SDev**
QuarterlyAverageBalance	4,349	87,865	24,288	287,216	78,455	834,951
NumberOfCredits	7	38	21	44	47	298
NumberOfDebits	11	34	38	75	65	250
TotalRisk	42,728	369,137	198,405	838,516	642,365	1,581,504
TotalLimit	163,107	989,558	722,198	1,959,832	2,554,236	4,062,079
TotalRisk	666,718	17,944,666	968,002	4,016,337	10,145,128	28,177,464
TotalLimit_IOB	1,915,748	44,171,101	3,234,133	17,237,407	30,439,247	87,045,370
CashRiskRatio	0.2187	0.4118	0.3923	0.4741	0.4486	0.4778
CentralBankPercentageOfChequesReturned	0.0307	0.0251	0.0379	0.0204	0.0382	0.0202
PercentageOfChequesReturned	0.0040	0.0517	0.0113	0.0758	0.0192	0.0867
LimitUtilizationRatio	0.2185	0.3706	0.2935	0.3333	0.2655	0.3057
LimitUtilizationRatio_IOB	0.2195	2.4587	0.6030	17.4863	0.4522	3.3651
Age_at_Loan	107.5472	73.8261	113.2237	79.1459	149.7536	85.1148
Check_Length	2.6462	5.9414	3.5758	8.7528	4.0494	10.2312
NumberOfBanks	1.5887	2.8784	4.6729	4.2241	10.4589	6.3658
ColletralType1	104	3,128	2,964	47,850	5,549	68,693
ColletralType2	43,961	497,813	258,655	1,804,986	567,839	2,705,453
ColletralType3	7,876	31,341	14,318	175,584	42,802	859,126
ColletralType4	238,142	2,164,538	1,403,832	4,429,936	3,864,050	6,992,758

Table 5 : Summary statistics of time variant variables

	Small Firms	Medium Firms	Large Firms	Full Sample
Number of Borrowers	3,667	8,602	3,324	15,593
Number of Observations	49,047	107,858	45,710	202,615
Event	595	549	163	1,307
Export	4	25	61	90
Import	191	1,309	947	2,447
Unknown Sector	2,355	551	45	2,951
Fishing Sector	1	12	6	19
Social Services Sector	78	251	48	377
Education Sector	9	79	19	107
Electricity, Gas and Water Resources Sector	94	141	57	292
Real Estate Trading Sector	17	249	63	329
Financial Intermediation Sector	14	44	8	66
Manufacturing Sector	307	2,320	1,335	3,962
Construction Sector	286	1,834	358	2,478
Sole Proprietorship Sector	10	60	13	83
Mining and Quarrying Sector	6	80	35	121
Hotels and Restaurants Sector	37	185	32	254
Wholesale and Retail Trade Sector	244	1,759	1,015	3,018
Defense, Public Administration and Social Security Sector	1	4	5	10
Agriculture, Hunting and Forestry Sector	33	269	118	420
Transportation, Warehousing and Communications Sector	36	370	91	497
Small Enterprises Sector	84	63	2	149
Health and Social Services Sector	55	331	74	460
Akdeniz Region	266	655	279	1,200
Doğu Anadolu Region	170	521	69	760
Ege Region	202	608	308	1,118
Güneydoğu Anadolu Region	122	558	160	840
İç Anadolu Region	704	1,647	504	2,855
Karadeniz Region	241	786	202	1,229
Marmara Region	1,962	3,827	1,802	7,591

Table 6 : Summary statistics of time invariant variables

5. Empirical Analysis

5.1. Objective

In this study we are trying to forecast the financial distress of SMEs and corporate enterprises, utilizing survival analysis methodology. Financial distress prediction is commonly used in banking especially in developed countries, since the early 1970'es[22]. Such practices have started to become prevalent in Turkey only recently. Although the best banks were already using such methods, for many others the need to comply with recently enacted[23] regulatory requirements of BRSA has been the main spur.

5.2. Methodology

Although bankruptcy happens at a point in time, it is not an isolated incident, it is the end result of failure processes occurring gradually over a period of time. Therefore, considering the time related information is a very important input in modeling financial distress. This means the model should be flexible enough to incorporate firm level and macro variables that vary with calendar time.

Survival analysis has been developed to handle such considerations. It has three important features distinguish it from the earlier dichotomous methods discussed in Chapter 3.

---

[22] Altman's model that we have already mentioned in Chapter 3, has been commercially implemented since the 1970'es and is still the most commonly used technique in financial distress prediction.

[23] Credit Risk Calculation based on Internal Rating Approach (2014) (Kredi Riskine Esas Tutarın İçsel Derecelendirmeye Dayalı Yaklaşımlar ile Hesaplanmasına İlişkin Tebliğ)

i. Unlike the dichotomous approach of failure vs survival which is per force static and has to arbitrarily deal with a single point in time, survival analysis is dynamic and deals with time until failure

ii. Survival analysis can thus incorporate time-varying covariates i.e. it can examine the time series path of how the firms performs by using these covariates. On the other hand, the dichotomous methods can analyze only incorporate variables pertaining to a single point in time.

iii. The presence of censored observations provides extra information. Censoring occurs if the survival time of the firm is longer than the observation period or the firms are excluded from observations for reasons other rather than financial distress.

Among the models based on survival analysis, Cox (1972)'s proportional hazard model has been the most widely used in various fields. Its popularity is due to its being free of distributional assumptions which makes it robust. A good survey of survival analysis from an applied accounting and finance perspective is contained in LeClere (2008).

5.2.1. Survival Analysis

Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs (Kleinbaum, 1996). T**ime**, can be days, weeks, months or quarters etc. from the beginning of analysis or follow-up until an event occurs. By **event** we mean, death, recovery, bankruptcy or any designated experience of interest that may happen to the entity being tracked. The entity can be person, firm etc.

Although more than one event may be considered in the same analysis we assume that the entity experiences only one event. A key analytical problem called censoring occurs when we don't know the survival time exactly. There generally three reasons why censoring may occur;

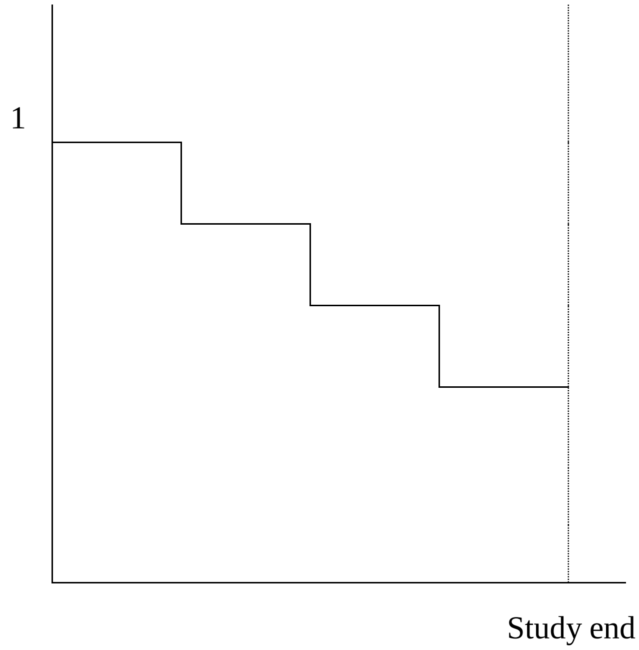
i. An individual doesn't experience the event before the study ends,
ii. An individual is lost to follow-up during the study period,
iii. An individual withdraws from the study because of an event.

Hazard and survival functions are two main important concepts in survival analysis. The survival function, denoted by *S(t)* is the probability that an individual survives longer than some specific time *t*.

$$s(t) = P(T \geq t)$$

$$S(t) = P(T \geq t) = 1 - P(T \leq t) = 1 - F(t)$$

Where $f(t) = F'(t)$. Here f(t) means probability density and F(t) means cumulative density or distribution function respectively. Survival functions are usually step functions illustrated as follows.



The hazard function is defined as the limit of the conditional probability

$$h(t) = \lim_{\Delta t \to 0^+} \frac{P(t < T < t + \Delta t | T > t)}{\Delta t}$$

which gives the instantaneous potential per unit time for the event to occur, given that the individual or entity has survived up to time *t*. The hazard function *h(t)* means the potential for failure at time *t* per unit time. Also,

$$h(t) = \frac{f(t)}{S(t)} = \frac{S'(t)}{S(t)}$$

Survival functions *S(t)*, hazard rate *h(t)* and the cumulative density or distribution function *F(t)* gives the same information.

The fundamental quantity used to access the risk of event occurrence in each discrete time period is known as hazard. Discrete time hazard *is the conditional probability that* *individual i will experience the event in time period j, given that he or she did not experience it in an earlier time period.*

Although it is not directly related with our study we would like to give information about findings which may help bank to develop its screening or lending process capabilities. These findings may also help the bank to evaluate the impact of major changes in lending and monitoring processes.

		Number of			Proportion of	
Quarter	Time Interval	# of Customers at the beginning of the quarter	# of Defaulters during the quarter	Censored at the end of the quarter	Customers at the beginning of the quarter who defaulted during the quarter	All customers still active at the end of the quarter
0	[0,1)	15,593	-	-	-	1.0000
1	[1,2)	15,593	105	983	0.0067	0.9933
2	[2,3)	14,505	36	394	0.0025	0.9908
3	[3,4)	14,075	81	681	0.0058	0.9850
4	[4,5)	13,313	128	644	0.0096	0.9754
5	[5,6)	12,541	118	705	0.0094	0.9660
6	[6,7)	11,718	111	503	0.0095	0.9565
7	[7,8)	11,104	105	597	0.0095	0.9471
8	[8,9)	10,402	95	605	0.0091	0.9379
9	[9,10)	9,702	92	616	0.0095	0.9285
10	[10,11)	8,994	74	485	0.0082	0.9202
11	[11,12)	8,435	49	563	0.0058	0.9144
12	[12,13)	7,823	50	622	0.0064	0.9080
13	[13,14)	7,151	53	766	0.0074	0.9006
14	[14,15)	6,332	28	612	0.0044	0.8962
15	[15,16)	5,692	29	522	0.0051	0.8911
16	[16,17)	5,141	17	344	0.0033	0.8878
17	[17,18)	4,780	26	366	0.0054	0.8824
18	[18,19)	4,388	14	325	0.0032	0.8792
19	[19,20)	4,049	11	365	0.0027	0.8765
20	[20,21)	3,673	19	330	0.0052	0.8713
21	[21,22)	3,324	10	301	0.0030	0.8683
22	[22,23)	3,013	9	247	0.0030	0.8653
23	[23,24)	2,757	7	285	0.0025	0.8627
24	[24,25)	2,465	6	220	0.0024	0.8603
25	[25,26)	2,239	12	196	0.0054	0.8550
26	[26,27)	2,031	3	186	0.0015	0.8535
27	[27,28)	1,842	3	202	0.0016	0.8518
28	[28,29)	1,637	6	182	0.0037	0.8482
29	[29,30)	1,449	2	162	0.0014	0.8468
30	[30,31)	1,285	4	181	0.0031	0.8437
31	[31,32)	1,100	4	1034	0.0036	0.8401
32	[32,33)	62	0	62	-	0.8401

Table 7 : Life table

Risk Set

Hazard function

Survivor Function

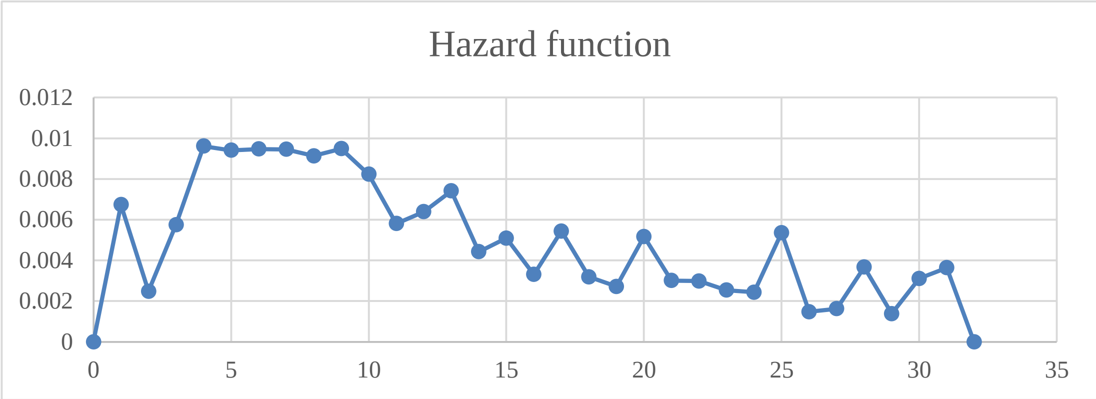


Figure 8 : Hazard function

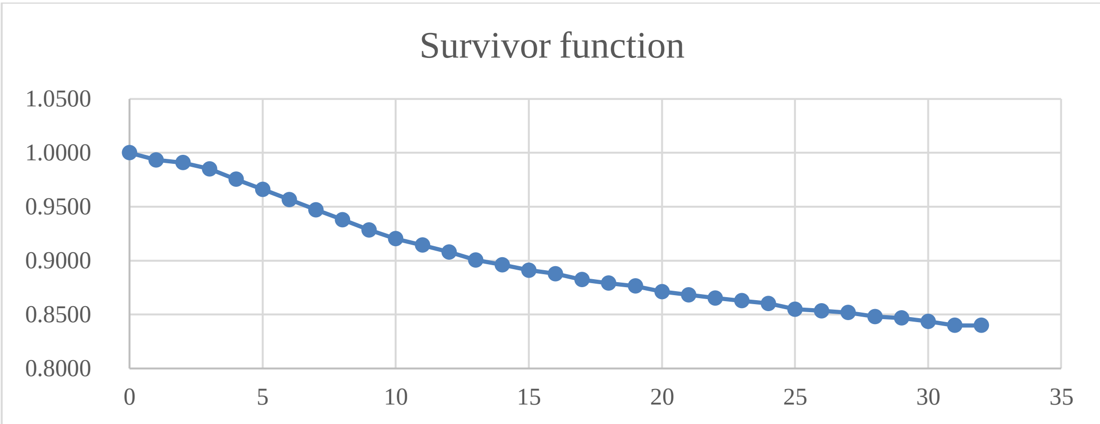


Figure 9 : Survivor function

It is very important to note that;

- ✓ 8.03 % (105 defaulters out of 1.307 from Table 7 whose quarter is equal to 1) of all defaulters experience financial distress i.e., default within the first quarter just after their credit is granted.
- ✓ 96,2 % (101 out of 105 defaulters) of those which default within the first quarter are small firms as classified by their sales, i.e., dSales_s = 1
- ✓ Bank should improve screening capabilities for especially small firms

5.2.2. The Cox PH-Proportional Hazard and its Characteristics

The Cox PH model is a mathematical model for analyzing survival data and it is usually written in terms of the hazard formula

$$h(t, X) = h_0(t)e^{\sum_{i=1}^{p} \beta_i X_i}$$

Where $X = (X_1, X_2, X_3, \dots \dots . X_p)$ explanatory or predictor variables. This model gives an expression for the hazard at time *t* for an individual with a given specification of a set of explanatory variables denoted by *X*.

The Cox model formula says that the hazard at time t is the product of two quantities.

The first of these, $h_0(t)$ is called the baseline hazard function. The Cox model formula has the property that if all the X's are equal to zero, the formula reduces to the baseline hazard function. That is, the exponential part of the formula becomes e to the zero, which is 1. This property of the Cox model is the reason why $h_0(t)$ is called the **baseline function**.

$$h(t, X) = h_0(t)e^{\sum_{i=1}^{p} \beta_i X_i} \; where \; X_1 = X_2 = \cdots .. = X_p = 0$$

$$= h_0(t)e^0$$

$$= h_0(t)$$

From a slightly different point of view, the Cox model reduces to the baseline hazard when no X's are in the model. Thus $h_0(t)$ may be considered as a starting or "baseline" version of the hazard function, prior to considering any of the X's.

The second quantity is the exponential expression e to the linear sum of $\beta_i X_i$, where the sum is over the p explanatory X variables. It involves X's but not t's. The X's here are called **time-invariant** X's. The region, sector, date of foundation of a firm or gender of an individual are examples of time-invariant attributes we call X.

It is possible that X's may involve t. Such X's are called **time-varying** X's. Quarterly average balance, limit and risk values of a firm or monthly income and weight of an individual are examples of time varying attributes we call X.

Another important property of the Cox model is that the baseline hazard, $h_0(t)$ is an unspecified function. It is this property that makes the Cox model a **nonparametric** model. [24]

A key reason for the popularity of the Cox model is that, even though the baseline is not specified, reasonably good estimates of regression coefficients and hazard ratios of interest, and adjusted survival curves can be obtained for a wide variety of data situations. Another way of saying this is that the Cox PH model is a robust model, so that the results from using the Cox will closely approximate the results for the correct but unknown parametric model.

[24] As the term nonparametric implies, parametric models are also possible. In such models one replaces the unspecified h(t) with a fully specified function like the log normal or Weibull. Such models are beyond the scope of our study.

In general, a hazard ratio *(HR)* is defined as the hazard for one individual or entity divided by the hazard for a different individual or entity. The two entities being compared can be distinguished by their values for the set of predictors, that is, the X's.

We can write the hazard ratio as the estimate of $h(t, X^*)$ divided by the estimate of $h(t, X)$, where X^*denotes the set of predictors for one individual, and X denotes the set of predictors for the other individual.

$$\widehat{HR} = \frac{\hat{h}(t, X^*)}{\hat{h}(t, X)} = \frac{\hat{h}_0(t)e^{\sum_{i=1}^{p} \hat{\beta}_i X_i^*}}{\hat{h}_0(t)e^{\sum_{i=1}^{p} \hat{\beta}_i X_i}}$$

$$\widehat{HR} = \ e^{\sum_{i=1}^{p} \beta_i (X_i^* - X_i)}$$

We now obtain an expression for the *HR* formula in terms of the regression coefficients by substituting the Cox model formula into the numerator and denominator of the hazard ratio expression. This substitution is shown here. Note that the only difference in the numerator and denominator are the X^*'s versus the X 's. Notice also that the baseline hazards will cancel out.

Using algebra involving exponentials, the hazard ratio formula simplifies to the exponential expression shown here. Thus, the hazard ratio is computed by exponentiating the sum of each β_i “hat”times the difference between $X_i^*\ and\ X_i$

The PH assumption requires that the *HR* is constant over time, or equivalently, that the hazard for one individual is proportional to the hazard for any other individual, where the proportionality constant is independent of time.

A graphical way of testing the PH assumption, involves reconsidering the HR formula comparing two different entities or individuals.

By comparing Log-log Survival (LLS) curves we can test whether the proportionality assumption holds. Basically one checks whether the estimated LLS curves are parallel to each other. The general formula for Kaplan Meier (KM) estimate of the survival function *S(t)* is

$$\hat{S}(t_f) = \prod_{i=1}^{f} \widehat{Pr}\{T > t_i | T \geq t_i\}$$

The actual survival probability can be plotted using KM method. All LLS curve is simply a transformation of an estimated survival curve taking log twice.

$$-ln[-ln\, S\,(t, X)] = \; -ln\left[-ln\left(S_0(t)^{e^{\sum_{i=1}^{p} \beta_i X_i}}\right)\right]$$

$$= -ln\left[-e^{\sum_{i=1}^{p} \beta_i X_i} * lnS_0(t)\right]$$

$$= \sum_{i=1}^{p} \beta_i X_i - ln[-lnS_0(t)]$$

Because two different individuals X_1 and X_2 have two different LLS formulas,

$$-ln[-ln\, S(t, X_1)] = \sum_{i=1}^{p} \beta_i X_{1i} - ln[-ln\, S_0(t)]$$

$$-ln[-ln\, S(t, X_2)] = \sum_{i=1}^{p} \beta_i X_{2i} - ln[-ln\, S_0(t)]$$

The expression for subtracting one from the other does not depend on time, but constant as follows:

$$-ln[-lnS(t, X_1)] - ln[-lnS(t, X_2)] = \sum_{i=1}^{p} \beta_i (X_{1i} - X_{2i})$$

Because the difference, $\sum_{i=1}^{p} \beta_i (X_{1i} - X_{2i})$ does not involve time t, the estimated LLS curves would be approximately parallel under the proportionality assumption. (Nam, Kim, Park, & Lee, 2008).

We performed the above visual test when firms are categorized by Limit Utilization Ratio, Number of Staff, Construction Sector and Import. These graphs are presented in Subsection 5.4.4 as Figure 16, Figure 17, Figure 18 and Figure 19. As will be seen for the first three categories, the lines intersect and thus indicate a violation of the proportionality assumption. In other words there is a time pattern to failures when firms are categorized by, say, Limit Utilization Ratio. But Import's impact on failure is independent of time i.e. the graphs for Import=0 and Import=1 do not intersect, see Figure 19.

The above graphical test is intuitively appealing but can only be used with categorical covariates or explanatory variables[25]. There is also a general statistical method of checking the PH assumption. The `estat ptest` command in STATA tests for individual covariates and globally the null hypothesis of zero slope, which is equivalent to testing that the log hazard ratio function is constant over time. Thus rejection of the null hypothesis of a zero slope indicates deviation from the proportional-hazards assumption. (http://www.stata.com/manuals13/ststcoxph-assumptiontests.pdf)

[25] We will be discussing the test of proportionality assumption in Subsection 5.4.2.

5.3. Data Analysis and Model Estimation

We started by checking the integrity of the data based on knowledge of banking, i.e. total limit cannot be greater than total limit for each customer during each period, or date of foundation should be smaller than the date of granting credit and the like. Thus, we excluded such outlier observations from the data set.

As mentioned before, a key feature of the proportional hazards model is that the model can utilize time-series variations in the covariates. The model can provide information on whether changes in explanatory variable over time influence the probability of the event occurring.

Using the input variables already presented in Section 4.2, we estimated our Cox PH model as discussed in Subsection 5.2.2 . This model is referred to as *Model 1*. As mentioned our original panel data set is from 2005 to 2012. We define this as *[2005, 2012] window*.

To test the model's predictive performance in different time periods, we have used window data i.e.,

i. We have deleted the observations belonging to the year 2012 so that the panel data ended in 2010. We called this as *[2005, 2011] window*.
ii. We have deleted the observations belonging to the year 2011 and 2012 so that the panel data ended in 2011. We called this as *[2005, 2010] window*.

The Cox proportional hazard model estimates a relative Risk Score[26] for each firm. This score is relative to a baseline group. The baseline group is constant and identifiable when data is time invariant_like 30 year old males with a prior delinquency. Unfortunately with time varying data such as ours, each firm's reference group changes in ways that the modeler can not control. When the Risk Score is greater than 1 the firm in question has a greater rate of undergoing failure per quarter than its reference group. Thus we flagged such firms as potential defaulter.

We have also used RiskScoreGreaterThan1 ratio for our analysis but after a few executions we realized that this ratio does not provide any additional predictive capability, and to avoid complexity, we did not show analysis results in our study.

Finally to calibrate the model (i.e. minimize selectivity ratio[27] for defaulters) we also used scales 2, 3, 4 and 5 to flag the customers as potential defaulter and compared with the actual defaulters. We call this Calibration.

To evaluate *Model 1*'s predictive performance, we have compared actual versus predicted defaults

- Within sample prediction,

---

[26] These risk scores are very closely related to the predicted hazard for each firm, Singer & Willett (2003) p. 535-542. We provide some graphs of estimated hazards in Subsection 5.4.4.

[27] Selectivity ratio is the proportion of #of flagged customer to #of all active and non-default customers at the last quarter of the window.

- Out of sample prediction, namely, 1 quarter ahead and 1 year ahead prediction of the actuals,

With three[28] different windows and Calibration approach. We have showed analysis results belonging to Model 1 in Subsection 5.4.1, Model 2 in Section 7.1 and Section 7.2.

In Model 1, we have ignored the possibility of assumptions violation, particularly by our time varying covariates. We identified the variables violating the proportionality assumption by using `estat phtest` command in STATA and found that the variables "*number of staff quarterly*" and *Limit utilization ratio*" violated proportionality assumption in Subsection 5.4.2. Intuitively these variables impact on risk were changing across time periods instead of being constant as assumed by the Cox model.

To fix the violation of proportionality assumption, we used two methods:

i. TVC[29] command in STATA for utilizing time-varying covariates. We called this as *Model 2*.

ii. Breaking Points method. We called this as *Model 3*.

We explain the details in Subsection 5.4.2.

Like we did for Model 1, we evaluated *Model 2'*s predictive performance in Section 7.1 and *Model 3's* predictive performance in Section 7.2.

---

[28] Namely, [2005, 2012] window, [2005, 2011] window and [2005, 2010] window

[29] The TVC is used to specify those variables that vary continuously with respect to time, i.e., time-varying covariates.

We then compared the predictive performances of Model 1, Model 2 and Model 3 in Subsection 5.4.5.

Thus in total we have three models and three time windows. This means we have performed 9 separate Cox regressions and estimated their hazard ratios.

We have summarized other findings that in Subsection 5.4.4.

Finally, we have tried to simulate what if the bank had used its own internal rating system for financial distress prediction and compared to our findings in Subsection 5.4.5.

5.4. Results of the Analysis

5.4.1. Model 1

5.4.1.1. [2005, 2012] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9447692	0.727
Ege Region	1.209667	0.17
Güneydoğu Anadolu Region	0.6777609	0.05
İç Anadolu Region	0.9899246	0.933
Karadeniz Region	1.002555	0.986
Marmara Region	1.039013	0.722
Fishing Sector	1.56597	0.528
Social Services Sector	0.73743	0.133
Electricity, Gas and Water Resources Sector	0.3803172	0.011
Real Estate Trading Sector	0.3613052	0.007
Education Sector	0.2834308	0.075
Financial Intermediation Sector	0.239475	0.154
Mining and Quarrying Sector	0.7221758	0.392
Hotels and Restaurants Sector	0.6032622	0.132
Defense, Public Administration and Social Security Sector	3.46101	0.081
Health and Social Services Sector	0.43132	0.001
import	0.652553	0
export	0.4010897	0.362
dsales_l	0.3948513	0
dsales_m	0.4679868	0
numberofstaffquarterly	0.8921726	0
quarterlyaveragebalance	0.9999458	0
numberofcredits	0.8582893	0
numberofdebits	1.002458	0.273
totalrisk	0.9999998	0.052
totallimit	1	0.39
limitutilizationratio	1.019373	0.843
cashriskratio	4.008199	0
totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.001153	0.232
age_at_loan	0.9965147	0
check_length	1.023257	0

numberofbanks	1.093668	0
rtrnedchecksprctn	5.810139	0
cb_returnedchecksprcnt	719924.7	0
w_collateral_1	0.2729056	0.442
w_collateral_3	1.027097	0.871
w_collateral_2	5.299396	0

Cox regression -- Breslow method for ties

No. of subjects =	15593	Number of obs =	202615
No. of failures =	1307		
Time at risk =	202615		
		LR chi2(39) =	2673.37
Log likelihood =	-10636.664	Prob > chi2 =	0.0000

Table 8 : Cox regression with [2005, 2012] window for Model 1

Table 8 gives us incredible information about the borrower's tendency to potential default. Sector_14 - the defense, public administration and social security sector seem to be the most problematic sector. Sector_05 - electricity, gas and water resources, sector_06 - real estate trading and sector_18 - health and social services can also be valuable for better screening. Note that all three of them have hazard ratios seriously below 1. They have significant z values (absolute value of z is greater than 1.96)

Firms working in goods' import have a potentially less probability of default. These firms are more likely to hire more qualified staff which may lead a more professional management capability.

Normally we would expect low financial distress problem as borrowers' check_length value gets bigger but the very significantly POSITIVE coefficient of check_length suggests "**bad screening practices**" by Bank.

Variables which have significantly NEGATIVE coefficients (like Import and age_at_loan etc.) can also be valuable for better screening.

A %1 increase in Cash risk ratio means, the borrower is four times more likely to experience default compared to the baseline function.

Rtrnedchecksprctn[30] is an important variable for the monitoring activities. Like cash risk ratio, a %1 increase in Rtrnedchecksprctn means, the borrower is 5,8 times more likely to experience default compared to the baseline function.

The "w_Collateral_1[31] to w_Collateral_4" coefficients could also be useful to bank in deciding what kind of collateral to accept.

Note that these findings are also consistent with Model 2 and Model 3 including all time windows. Hence in order to avoid repetition we did not mention these findings in the explanation part of Model 2 and Model 3.

[30] Ratio of sum of bounced checks to total checks sent to ICH in a given quarter for a given customer

[31] Ratio of CollateralType1 to the sum of CollateralType1 to CollateralType4

Event		0	1	Total
0	Count	10,053	4,233	14,286
	Row %	70.37	29.63	100.00
	Column %	97.76	79.72	91.62
	Total %	64.47	27.15	91.62
1	Count	230	1,077	1,307
	Row %	17.60	82.40	100.00
	Column %	2.24	20.28	8.38
	Total %	1.48	6.91	8.38
Total	Count	10,283	5,310	15,593
	Row %	65.95	34.05	100.00
	Column %	100.00	100.00	100.00
	Total %	65.95	34.05	100.00

Table 9 : Actual vs within sample prediction with [2005, 2012] window for Model 1

Before we start the analysis we should clarify the main idea illustrated in Table 9, a representation technique that is used in the upcoming sections. Columns represent the potential defaulters, while rows represent the actual defaulters.

Within the sample prediction is just an indication about how successfully the model's output and actual defaults overlap. It has no practical use for the bank.

It is clearly seen from the Table 9 that there are 15,593 customers subject to study and Model 1 has pre identified 1,077 out of 1,307 which were defaulted at different calendar time between 2005 to 2012. An important terminology here is the *Selectivity Ratio*, that is the proportion of number of flagged customers to number of customers. It is %34.05 (5,310 / 15,593) in this figure.

Another important terminology is *success rate for defaulters*, a ratio between 0 and 1, which gives the idea about the performance of the prediction model namely how well the model has predicted the actual defaults. In mathematical expression, it is ratio of number of actual defaults within the flagged customers to the number of actual defaults. It is %82.40 (1,077 / 1,373) in Table 9.

Actual Defaults in 2013Q1		0	1	Total
0	Count	8,064	2,458	10,522
	Row %	76.64	23.36	100.00
	Column %	99.76	98.44	99.45
	Total %	76.22	23.23	99.45
1	Count	19	39	58
	Row %	32.76	67.24	100.00
	Column %	0.24	1.56	0.55
	Total %	0.18	0.37	0.55
Total	Count	8,083	2,497	10,580
	Row %	76.40	23.60	100.00
	Column %	100.00	100.00	100.00
	Total %	76.40	23.60	100.00

Table 10 : Actual vs 1 quarter ahead prediction with [2005, 2012] window for Model 1

Table 10 is an important outcome of this study. This table shows us that there are 10,580 active customers by the end of 2012 i.e., the customer is in credit relationship with the bank (not right censored) and is not a defaulter. Note that the column values are generated at the end of 2012 whereas the row values are generated at the end of 2013 March. Q1 represents the first three months of the year namely January, February and March.

By using Model 1 and [2005, 2012] window data, Model 1 flags 2,497 customer as potential defaulter for 1 quarter ahead prediction i.e., for 2013 Q1. Remember that when the model is run we don’t know the actual defaulters yet for 2013.

This information is very valuable for the bank because now the bank has the list of potential defaulters for upcoming quarter and has the chance to take action. A clearly defined set of risk-mitigating actions can lead to significant reduction of exposure as we have already discussed in Section 2.9.

As time passes, by the end of 2013 Q1 we observe the actual the defaulters in 2013 in the first quarter. It is clearly seen that there are 58 customers who have experienced financial distress and defaulted. The beauty of the model is that our model has already predicted 39 out of 58 as potential defaulters by the end of 2012. The success rate for defaulters is %67.24 (39 / 58).

Actual Defaults in 2013		0	1	Total
0	Count	7,992	2,379	10,371
	Row %	77.06	22.94	100.00
	Column %	98.87	95.27	98.02
	Total %	75.54	22.49	98.02
1	Count	91	118	209
	Row %	43.54	56.46	100.00
	Column %	1.13	4.73	1.98
	Total %	0.86	1.12	1.98
Total	Count	8,083	2,497	10,580
	Row %	76.40	23.60	100.00
	Column %	100.00	100.00	100.00
	Total %	76.40	23.60	100.00

Table 11 : Actual vs 1 year ahead prediction with [2005, 2012] window for Model 1

As we are trying to predict further ahead in time, it is an expected result that the success rates for 1 year ahead prediction will be smaller than compared to 1 quarter ahead prediction. See Table 11.

The selectivity ratio remained the same since we have used the same flagging algorithm. On the other hand, success rate for defaulters decreased to %56.46 from %67.24 as expected. The model successfully predicted 118 customers out of 209 customers that defaulted in the year 2013.

5.4.1.2. [2005, 2011] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9783055	0.904
Ege Region	1.257787	0.136
Güneydoğu Anadolu Region	0.6693124	0.076
İç Anadolu Region	1.03486	0.796
Karadeniz Region	1.070759	0.659
Marmara Region	1.088786	0.477
Fishing Sector	1.099222	0.925
Social Services Sector	0.5736989	0.024
Electricity, Gas and Water Resources Sector	0.2867525	0.013
Real Estate Trading Sector	0.2751396	0.01
Education Sector	0.3551602	0.144
Financial Intermediation Sector	0.3202519	0.256
Mining and Quarrying Sector	0.6623592	0.36
Hotels and Restaurants Sector	0.4129917	0.049
Defense, Public Administration and Social Security Sector	3.704661	0.066
Health and Social Services Sector	0.4737732	0.008
import	0.6668413	0
export	0.5310177	0.528
dsales_l	0.3669269	0
dsales_m	0.4844538	0
numberofstaffquarterly	0.9156052	0
quarterlyaveragebalance	0.9998487	0
numberofcredits	0.8722336	0
numberofdebits	1.003324	0.197
totalrisk	0.9999998	0.106
totallimit	1	0.193
limitutilizationratio	0.8598964	0.165
cashriskratio	4.310642	0

totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.000824	0.572
age_at_loan	0.9961217	0
check_length	1.029114	0
numberofbanks	1.097208	0
rtrnedchecksprctn	5.530459	0
cb_returnedchecksprcnt	347100.5	0
w_collateral_1	0.0073666	0.312
w_collateral_3	1.183046	0.326
w_collateral_2	5.683613	0

Cox regression -- Breslow method for ties

No. of subjects =	13620	Number of obs =	161010
No. of failures =	1113		
Time at risk =	161010		
		LR chi2(39) =	2370.20
Log likelihood =	-8843.7054	Prob > chi2 =	0.0000

Table 12 : Cox regression with [2005, 2011] window for Model 1

Event		0	1	Total
0	Count	9,425	3,082	12,507
	Row %	75.36	24.64	100.00
	Column %	98.14	76.74	91.83
	Total %	69.20	22.63	91.83
1	Count	179	934	1,113
	Row %	16.08	83.92	100.00
	Column %	1.86	23.26	8.17
	Total %	1.31	6.86	8.17
Total	Count	9,604	4,016	13,620
	Row %	70.51	29.49	100.00
	Column %	100.00	100.00	100.00
	Total %	70.51	29.49	100.00

Table 13 : Actual vs within sample prediction with [2005, 2011] window for Model 1

Actual Defaults in 2012Q1		0	1	Total
0	Count	8,038	1,604	9,642
	Row %	83.36	16.64	100.00
	Column %	99.81	98.71	99.63
	Total %	83.05	16.57	99.63
1	Count	15	21	36
	Row %	41.67	58.33	100.00
	Column %	0.19	1.29	0.37
	Total %	0.15	0.22	0.37
Total	Count	8,053	1,625	9,678
	Row %	83.21	16.79	100.00
	Column %	100.00	100.00	100.00
	Total %	83.21	16.79	100.00

Table 14 : Actual vs 1 quarter ahead prediction with [2005, 2011] window for Model 1

Actual Defaults in 2012		0	1	Total
0	Count	7,956	1,566	9,522
	Row %	83.55	16.45	100.00
	Column %	98.80	96.37	98.39
	Total %	82.21	16.18	98.39
1	Count	97	59	156
	Row %	62.18	37.82	100.00
	Column %	1.20	3.63	1.61
	Total %	1.00	0.61	1.61
Total	Count	8,053	1,625	9,678
	Row %	83.21	16.79	100.00
	Column %	100.00	100.00	100.00
	Total %	83.21	16.79	100.00

Table 15 : Actual vs 1 year ahead prediction with [2005, 2011] window for Model 1

5.4.1.3. [2005, 2010] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9905926	0.962
Ege Region	1.214097	0.248
Güneydoğu Anadolu Region	0.7411445	0.207
İç Anadolu Region	0.9884942	0.936
Karadeniz Region	1.100354	0.563
Marmara Region	1.094233	0.484
Fishing Sector	1.314434	0.786
Social Services Sector	0.5754119	0.034
Electricity, Gas and Water Resources Sector	0.2849275	0.031
Real Estate Trading Sector	0.1807088	0.016
Education Sector	0.4251423	0.228
Financial Intermediation Sector	2.75E-20	.
Mining and Quarrying Sector	0.6547221	0.4
Hotels and Restaurants Sector	0.4179297	0.082
Defense, Public Administration and Social Security Sector	4.092786	0.048
Health and Social Services Sector	0.4150478	0.009
import	0.6857962	0.001
export	0.6287949	0.644
dsales_l	0.364825	0
dsales_m	0.5188525	0
numberofstaffquarterly	0.9326087	0
quarterlyaveragebalance	0.9997803	0
numberofcredits	0.8795584	0
numberofdebits	1.001477	0.646
totalrisk	0.9999998	0.158
totallimit	1	0.319
limitutilizationratio	0.7205445	0.007
cashriskratio	4.458559	0
totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.000759	0.627
age_at_loan	0.99602	0
check_length	1.031546	0
numberofbanks	1.106059	0
rtrnedchecksprctn	5.365207	0
cb_returnedchecksprcnt	124723.9	0
w_collateral_1	0.0005202	0.59
w_collateral_3	1.496571	0.024

w_collateral_2	6.492627	0

Cox regression -- Breslow method for ties

No. of subjects =	11804	Number of obs =	123758
No. of failures =	965		
Time at risk =	123758		
		LR chi2(38) =	2101.50
Log likelihood =	-7470.6032	Prob > chi2 =	0.0000

Table 16 : Cox regression with [2005, 2010] window for Model 1

Event		0	1	Total
0	Count	8,020	2,819	10,839
	Row %	73.99	26.01	100.00
	Column %	98.27	77.38	91.82
	Total %	67.94	23.88	91.82
1	Count	141	824	965
	Row %	14.61	85.39	100.00
	Column %	1.73	22.62	8.18
	Total %	1.19	6.98	8.18
Total	Count	8,161	3,643	11,804
	Row %	69.14	30.86	100.00
	Column %	100.00	100.00	100.00
	Total %	69.14	30.86	100.00

Table 17 : Actual vs within sample prediction with [2005, 2011] window for Model 1

Actual Defaults in 2011Q1		0	1	Total
0	Count	7,138	1,605	8,743
	Row %	81.64	18.36	100.00
	Column %	99.94	98.83	99.74
	Total %	81.43	18.31	99.74
1	Count	4	19	23
	Row %	17.39	82.61	100.00
	Column %	0.06	1.17	0.26
	Total %	0.05	0.22	0.26
Total	Count	7,142	1,624	8,766
	Row %	81.47	18.53	100.00
	Column %	100.00	100.00	100.00
	Total %	81.47	18.53	100.00

Table 18 : Actual vs 1 quarter ahead prediction with [2005, 2010] window for Model 1

Actual Defaults in 2011		0	1	Total
0	Count	7,082	1,565	8,647
	Row %	81.90	18.10	100.00
	Column %	99.16	96.37	98.64
	Total %	80.79	17.85	98.64
1	Count	60	59	119
	Row %	50.42	49.58	100.00
	Column %	0.84	3.63	1.36
	Total %	0.68	0.67	1.36
Total	Count	7,142	1,624	8,766
	Row %	81.47	18.53	100.00
	Column %	100.00	100.00	100.00
	Total %	81.47	18.53	100.00

Table 19 : Actual vs 1 year ahead prediction with [2005, 2010] window for Model 1

5.4.1.4. Calibration of Model 1

		Risk Score				
		>=1	**>=2**	**>=3**	**>=4**	**>=5**
[2005, 2012] window, Within Sample Prediction	**Selectivity Ratio**	34.05	15.94	9.85	6.91	5.37
	Success Rate	82.40	62.43	47.67	38.64	31.52
	Calibration Ratio	2.42	3.92	4.84	5.59	5.87
[2005, 2011] window, Within Sample Prediction	**Selectivity Ratio**	29.49	14.82	9.49	6.93	5.32
	Success Rate	83.92	65.77	52.47	42.50	35.31
	Calibration Ratio	2.85	4.44	5.53	6.13	6.63
[2005, 2010] window, Within Sample Prediction	**Selectivity Ratio**	30.86	16.03	10.58	7.77	6.11
	Success Rate	85.39	69.43	56.48	46.32	39.79
	Calibration Ratio	2.77	4.33	5.34	5.96	6.51
[2005, 2012] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	23.60	8.40	4.29	2.57	1.82
	Success Rate	67.24	44.83	34.48	24.14	15.52
	Calibration Ratio	2.85	5.33	8.04	9.39	8.51
[2005, 2011] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	16.79	6.22	3.31	2.14	1.32
	Success Rate	58.33	47.22	33.33	16.67	13.89
	Calibration Ratio	3.47	7.59	10.08	7.79	10.50
[2005, 2010] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	18.53	7.35	4.39	2.82	1.88
	Success Rate	82.61	56.52	39.13	21.74	4.35
	Calibration Ratio	4.46	7.69	8.91	7.72	2.31
[2005, 2012] window, 1Year ahead Prediction	**Selectivity Ratio**	23.60	8.40	4.29	2.57	1.82
	Success Rate	56.46	37.80	27.75	21.05	15.31
	Calibration Ratio	2.39	4.50	6.47	8.19	8.39
[2005, 2011] window, 1Year ahead Prediction	**Selectivity Ratio**	16.79	6.22	3.31	2.14	1.32
	Success Rate	37.82	25.00	16.03	10.26	8.33
	Calibration Ratio	2.25	4.02	4.85	4.80	6.30
[2005, 2010] window, 1Year ahead Prediction	**Selectivity Ratio**	18.53	7.35	4.39	2.82	1.88
	Success Rate	49.58	26.89	18.49	9.24	3.36
	Calibration Ratio	2.68	3.66	4.21	3.28	1.79

Table 20 : Calibration of Model 1

It is clear that as selectivity ratio is in the range between 0-1 and as it moves closer to 1 (i.e. flag most or all the customers as potential defaulter) the operational cost for monitoring activities gets higher. It is also worth noting that if selectivity ratio is 1 then success rate for defaulters will also 1 and will not provide any additional value to the bank. So the objective should be;

- Minimizing the selectivity ratio
- Maximizing the success rate for defaulters

The bank should decide the selectivity ratio range taking into consideration of operational cost per flagged customer, the size of the customer portfolio and a marginal decrease in success rate in predicting defaulters. The selectivity ratio in our model can be adjusted according to the bank's aspiration.

We have defined *calibration ratio* as success rate over selectivity ratio. Note that as risk score increases the selectivity ratio and success rate decrease simultaneously. As clearly seen from Table 20 that when selectivity ratio decreases to %8.40 from %23.60 the success rate for defaulters also decreases to %44.83 from %67.24 with [2005, 2012] window in Model 1.

The yellow highlighted cells are the greatest calibration ratios in a given row, it is to be noted that the greatest ratio are more likely to appear on the right side of the table.

This finding is also consistent with the calibration findings of the Model 2 and Model 3, we shall not detail more on this to avoid repetition.

5.4.2. Test of Proportionality Assumption

Time: Time			
Variable	**rho**	**chi2**	**Prob>chi2**
Doğu Anadolu Region	0.00559	0.04	0.8403
Ege Region	-0.04958	3.21	0.0732
Güneydoğu Anadolu Region	-0.01186	0.19	0.667
İç Anadolu Region	-0.0116	0.18	0.6749
Karadeniz Region	0.00012	0	0.9966
Marmara Region	-0.00219	0.01	0.9365
Fishing Sector	0.02106	0.58	0.4467
Social Services Sector	0.05978	4.68	0.0305
Electricity, Gas and Water Resources Sector	0.04013	2.15	0.1422
Real Estate Trading Sector	-0.00278	0.01	0.9198
Education Sector	-0.01606	0.34	0.5609
Financial Intermediation Sector	0.01809	0.43	0.5127
Mining and Quarrying Sector	-0.01811	0.43	0.5127
Hotels and Restaurants Sector	-0.00997	0.13	0.7184
Defense, Public Administration and Social Security Sector	-0.00962	0.12	0.7271
Health and Social Services Sector	0.16038	34.3	0
import	0.0375	1.86	0.1726
export	-0.0014	0	0.9595
dsales_l	0.14944	31.61	0
dsales_m	0.15292	36.79	0
numberofstaffquarterly	0.05343	64.2	0
quarterlyaveragebalance	0.00786	1.21	0.2711
numberofcredits	0.00694	0.28	0.5962
numberofdebits	-0.04174	1.92	0.1661
totalrisk	0.07871	7.23	0.0072
totallimit	0.06183	4.44	0.035
limitutilizationratio	0.22115	96.54	0
cashriskratio	0.02317	1.13	0.2873
totallimit_iob	-0.00572	0.12	0.7243
totalrisk_iob	-0.00045	0	0.9806
limitutilizationratio_iob	0.00432	0	0.9764
age_at_loan	0.05395	4.51	0.0336
check_length	-0.18521	45.59	0
numberofbanks	0.0239	1.01	0.316
rtrnedchecksprctn	-0.01859	0.4	0.5247
cb_returnedchecksprcnt	-0.06382	5.26	0.0219

w_collateral_1	0.00458	0.03	0.8741
w_collateral_3	-0.01273	0.22	0.6369
w_collateral_2	0.02483	1.1	0.295
global test		454.09	0

Table 21 : Test of proportionality assumption

We identified the variables violating the proportionality assumption by using estat phtest command in STATA and found that the variables *"Limit utilization ratio"* and *"number of staff quarterly"* violated proportionality. From the literature we know that, time varying covariates are more likely to violate the proportionality assumption compared to time invariant ones. That is why we did not dwell on "*Health and Social Services Sector, dsales_l and dsales_m"* although they too seem to violate the proportionality assumption as seen from the Table 21.

See also the graphical representations of Figure 16 for "limit utilization ratio" and Figure 17 for "number of staff quarterly". Notice the non-parallel or intersecting lines in these two graphs.

To fix the violation of proportionality and restore it, we used two methods, that is

iii. TVC command in STATA for utilizing time-varying covariates (Model 2)

iv. Breaking Points method. (Model 3)

In Breaking points technique we used business insight, and the central bank bounced cheques information. Business insight gave us the indication that, Turkish Banks began to feel the effects of the 2008 financial crisis by the 3rd quarter of 2008, and the bounced cheques ratio from the central bank showed that the crisis did affect the firms until the end of 2009.

Hence we splited *“limit utilization ratio”* and *“number of staff quarterly”* variables in terms of

- ✓ If date of observation < 2008Q3 → _precrisis,
- ✓ If date of observation >= 2008Q3 and date <= 2009Q4 → _crisis and
- ✓ If date of observation > 2009Q4 → _postcrisis

This is logical because the firms experience cash flow problems and generally decide downsizing. This is why limit utilization ratio and number of staff quarterly variables are more subject to change during crisis times.

5.4.3. Comparison Model 1, Model 2 and Model 3

	Model 1		Model 2		Model 3	
	Selectivity Ratio	Success Rate	Selectivity Ratio	Success Rate	Selectivity Ratio	Success Rate[32]
[2005, 2012] window within sample prediction	34.05	82.40	38.23	81.03	34.18	83.40
[2005, 2011] window within sample prediction	29.49	83.92	34.19	83.65	29.57	85.35
[2005, 2010] window within sample prediction	30.86	85.39	35.55	86.32	30.79	87.56
[2005, 2012] window 1 quarter ahead prediction	23.60	67.24	30.68	60.34	22.95	72.41
[2005, 2011] window 1 quarter ahead prediction	16.79	58.33	23.86	55.56	16.16	58.33
[2005, 2010] window 1 quarter ahead prediction	18.53	82.61	24.89	82.61	17.88	82.61
[2005, 2012] window 1 year ahead prediction	23.60	56.46	30.68	55.98	22.95	58.37
[2005, 2011] window 1 year ahead prediction	16.79	37.82	23.86	37.82	16.16	36.54
[2005, 2010] window 1 year ahead prediction	18.53	49.58	24.89	52.10	17.88	49.58

Table 48 : Performance comparison of Model 1, Model 2 and Model 3[33]

First, we consider the [2005, 2012] window with 1 quarter ahead prediction and try to compare performance of Model 1 and Model 2. We find a decrease in *success rate*[34] of Model 2 by %6.9 compared to success rate of Model 1, although we would normally expect an increase in the *success rate* in Model 2, as the selectivity ratio of Model 2 increases by %7.08 in compared to Model 1. The same logic applies to all windows

[32] Stands for “Success Rate for Defaulters”

[33] Selectivity ratios and success rates only represent risk scores for greater or equal to 1.

[34] Success rate stands for success rate for defaulters

including out of sample prediction and within sample prediction, so we can reach to a conclusion that performance of Model 1 is better than Model 2.

Predictability increases slightly in Model 3 compared to Model 1 hence we can say that Model 3 is the most powerful forecast model in our research.

Finally, breaking up the predictions by Default Year we get a predictable pattern. Namely our forecasting accuracy suffers as we move away from 2012. Namely since our model estimation uses 2005 to 2012 data, we get best results for 2013 followed by 2014 and the worst for 2015[35].

The model is stable for all windows – no fluctuations

5.4.4. Other Findings

Below we provide some graphs of estimated hazards for our sample. As discussed in Section 5.3, a firm's estimated hazard is very closely related to its risk score. Figure 10 presents these estimates when the firms in our sample are categorized by Sales size. As can be verified the large firms have the lowest hazard rate, followed by medium and small ones.

Figure 14 presents the estimated hazards when our firms are categorized by whether their Limit Utilization Ratio is lower or higher than their median. As can be observed, firms with larger Limit Utilization Ratio have a uniformly higher estimated hazard rate.

[35] The details of this finding is not given in the study to avoid complexity

Figure 15 does the same when our sample is categorized according to their geographic regions. Here the pattern is less clear cut than with Sales size. Firms in the Güneydoğu Anadolu region suffer the lowest hazard rate, prior to a very sharp increase in 2011, the ones in the Ege region have the highest hazard rate, the other regions are in between.

The model also estimates or predict a survival likelihood for each firm. These survival estimates are presented in Figure 10, Figure 11 and Figure 12. As can be expected hazard rate and survival likelihood are inversely related. In other words, a firm whose estimated rate of failure per quarter is smaller will have a larger likelihood of surviving. Figure 13, Figure 14 and Figure 15 conform to this expectation. Thus for instance firms with a lower than median Limit Utilization Ratio have a 99% chance of surviving until the 30'th quarter, whereas the comparable likelihood for a firm with higher than median Limit Utilization Ratio is 96%.

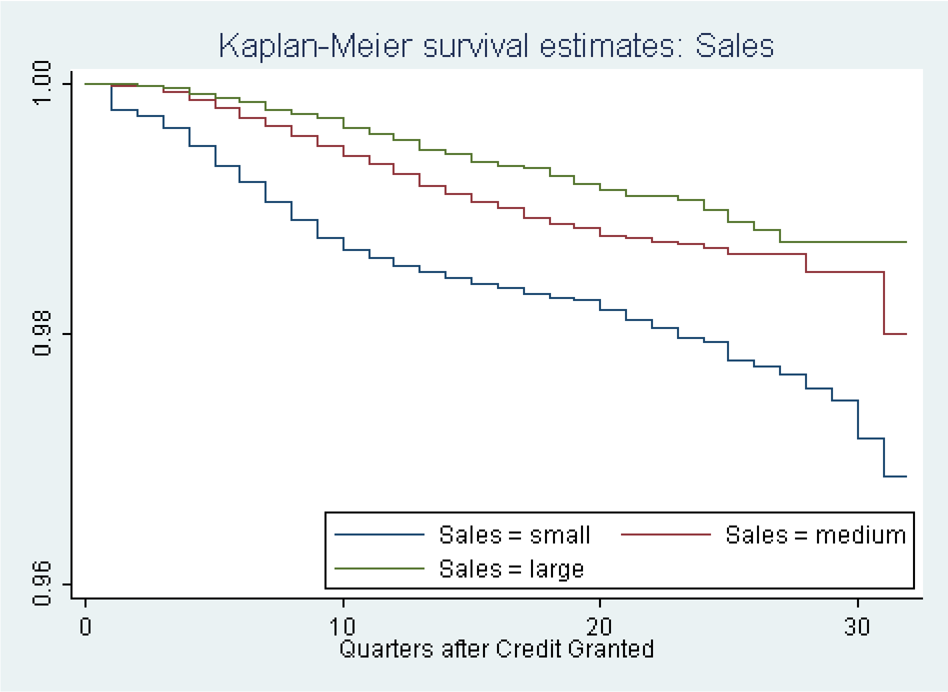


Figure 10 : Survival estimates of sales

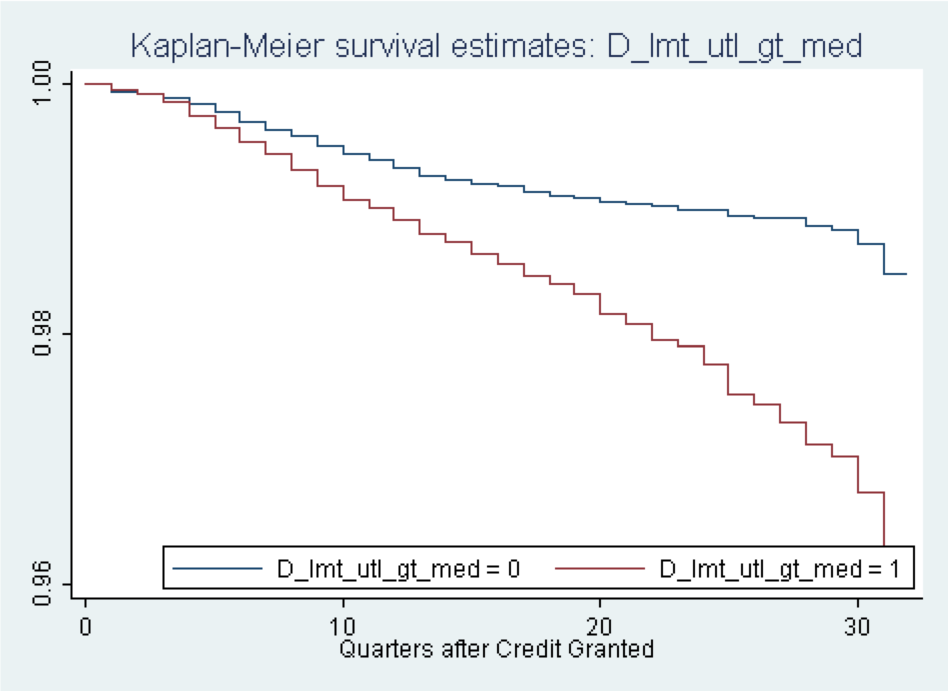


Figure 11 : Survival estimates of limit utilization ratio greater than median

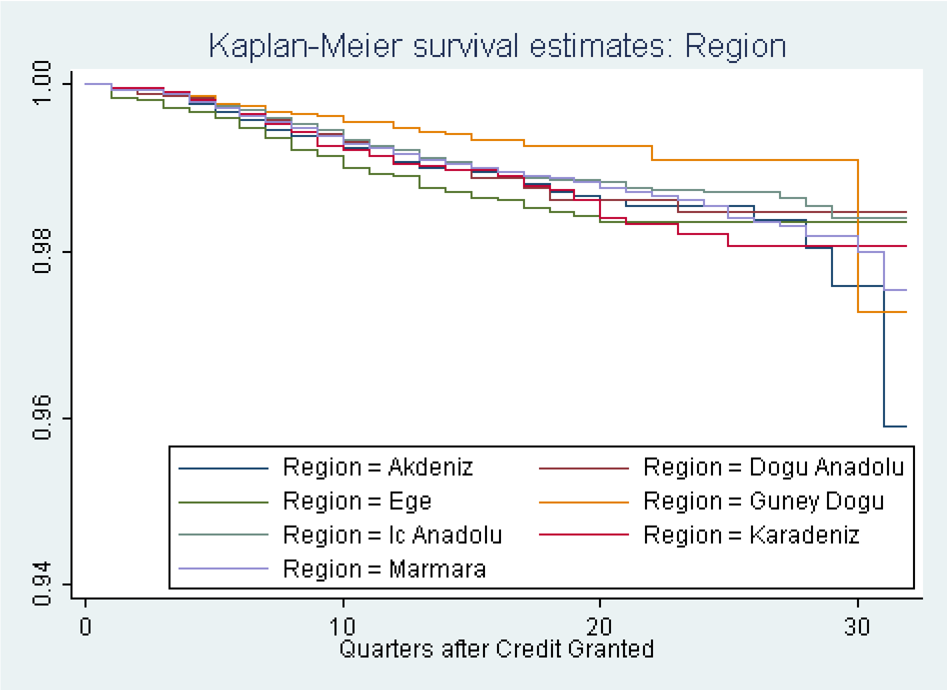


Figure 12 : Survival estimates of geographical regions

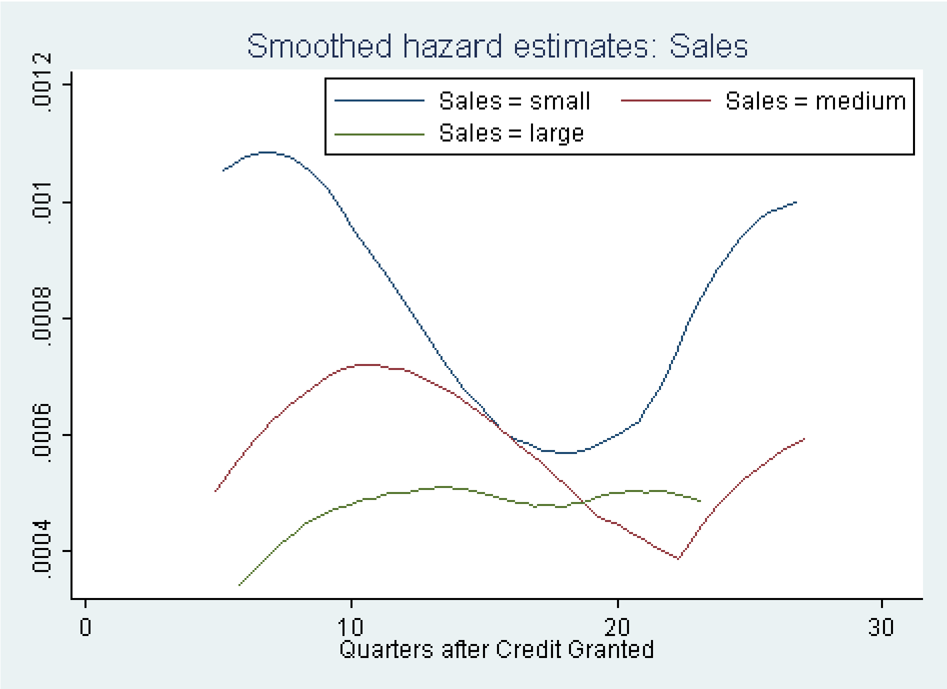


Figure 13 : Smoothed hazard estimates of sales

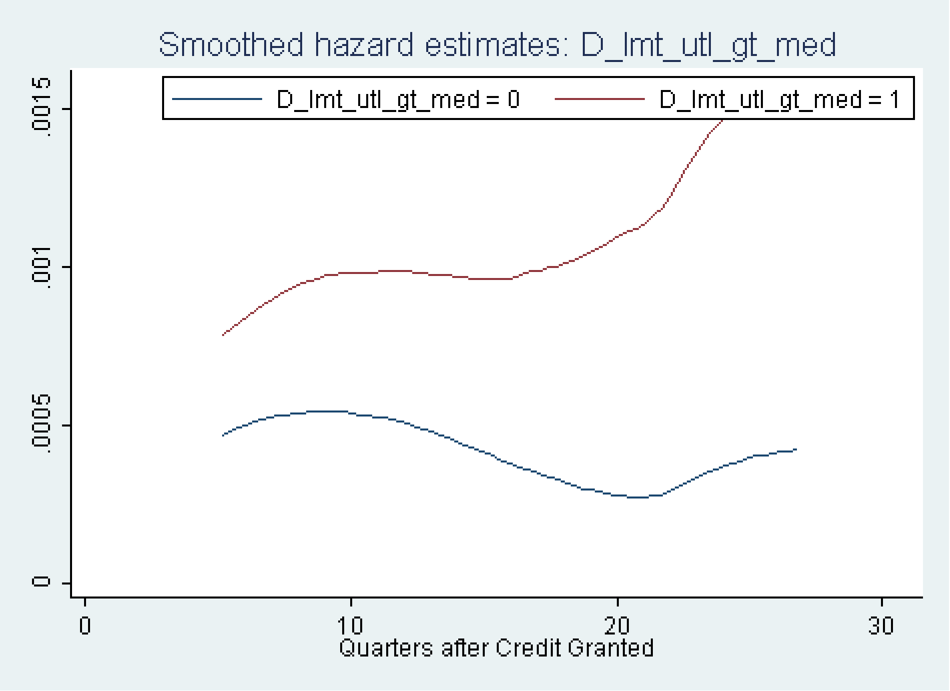


Figure 14 : Smoothed hazard estimates of limit utilization ratio

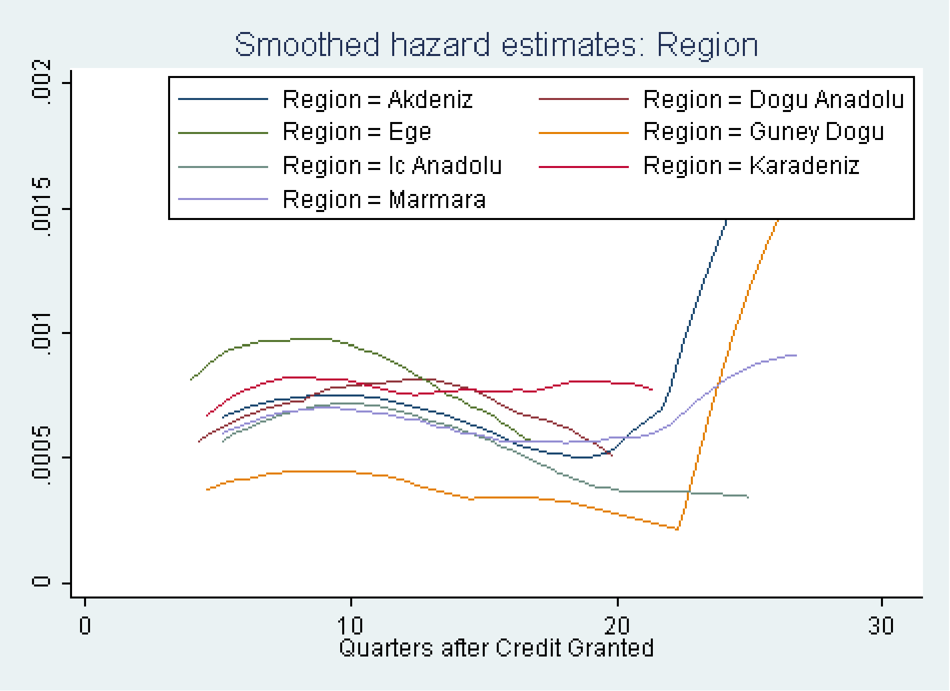


Figure 15 : Smoothed hazard estimates of geographical regions

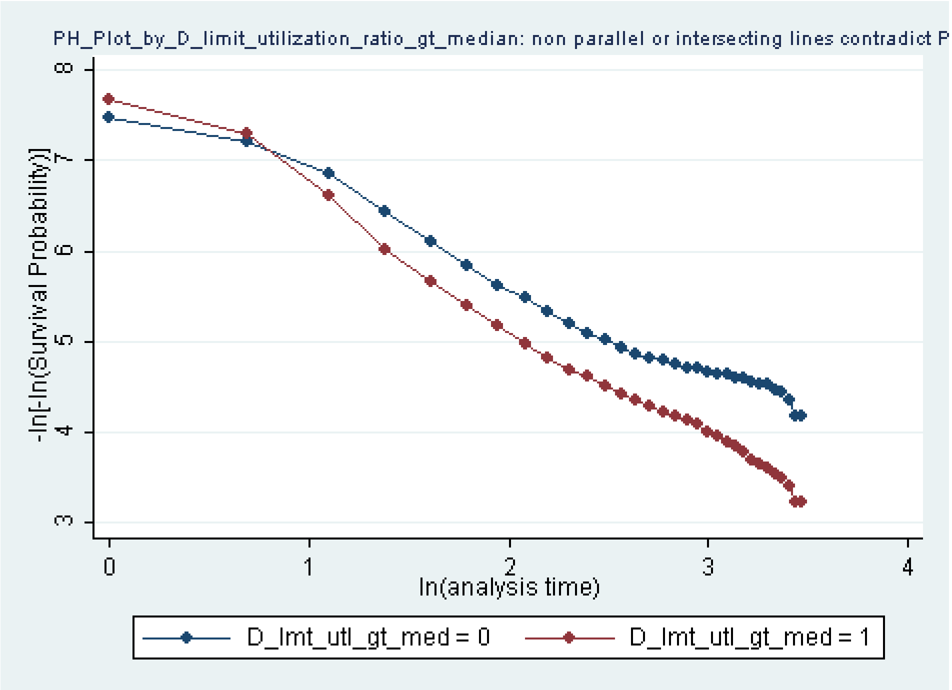


Figure 16 : Test of proportionality assumption of limit utilization ratio greater than median

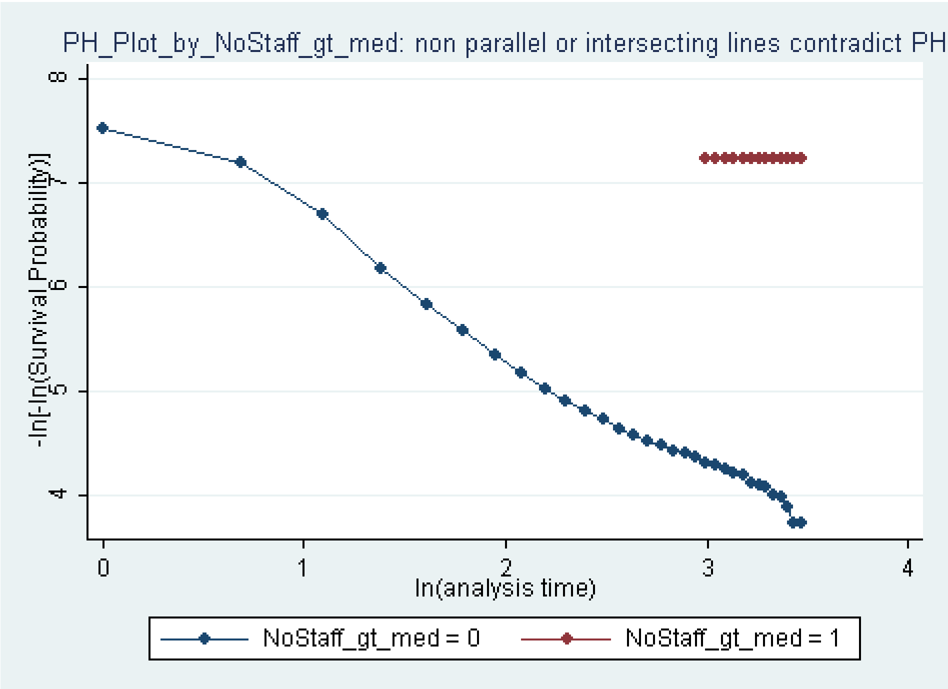


Figure 17 : Test of proportionality assumption of number of staff greater than median

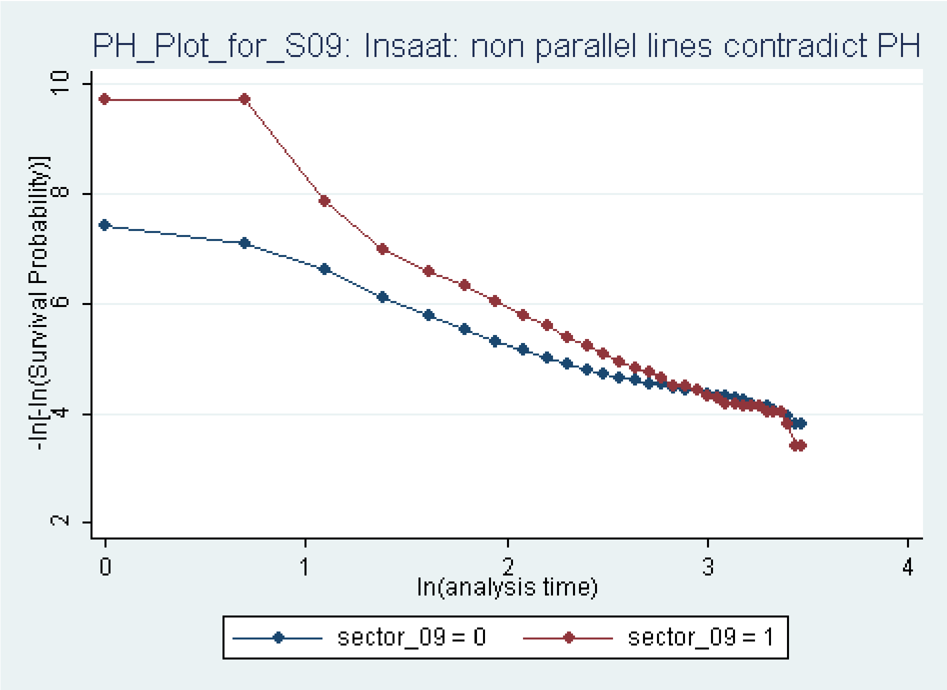


Figure 18 : Test of proportionality assumption of construction sector

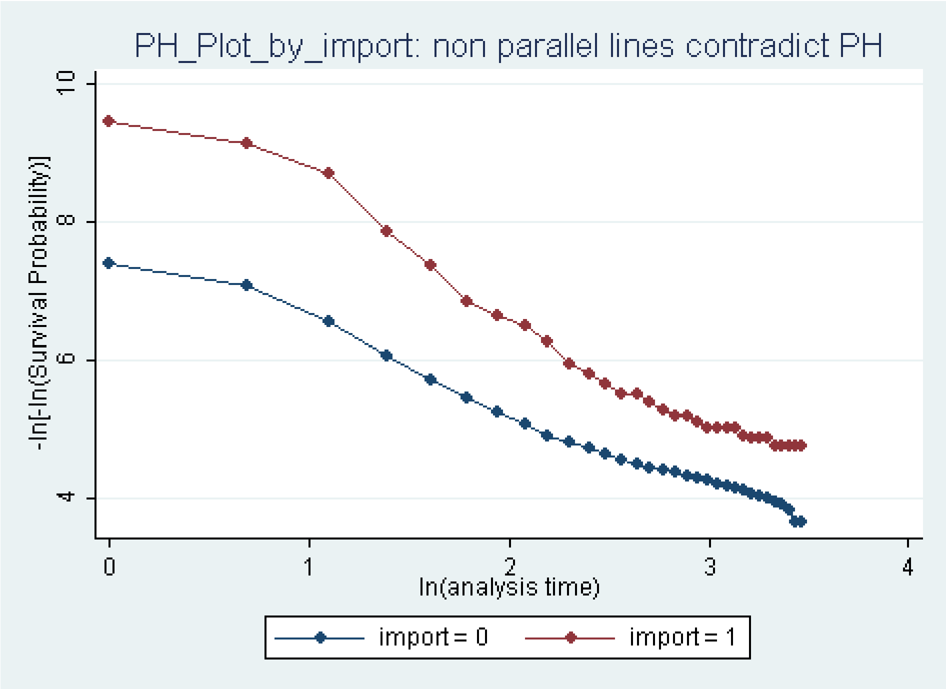


Figure 19 : Test of proportionality assumption of import

As we already mentioned in Section 2.11, BRSA have classified the collaterals according to their ability to convert cash in which the customer may able to pay back the loans. As we move from CollateralType1 to CollateralType4 the ability to convert to cash decreases, i.e. CollateralType1 and CollateralType2 are the strongest collateral that the bank may ask to borrower hence the cumulative ratio of these two types of collateral within all the collaterals gives us the hint about the risk appetite towards different customers from different geographical regions.

Ratio of Collateral Type1	Ratio of Collateral Type2	Cumulative Coll. Ratio	Akdeniz R.	Doğu Anadolu R.	Ege R.	Güneydoğu Anadolu R.	İç Anadolu R.	Karadeniz R.	Marmara R.
0.13%	8.66%	8.79%	0	0	0	1	0	0	0
0.11%	8.52%	8.63%	0	1	0	0	0	0	0
0.24%	8.00%	8.24%	0	0	0	0	1	0	0
0.41%	7.59%	7.99%	1	0	0	0	0	0	0
0.24%	7.67%	7.91%	0	0	0	0	0	1	0
0.12%	7.50%	7.62%	0	0	1	0	0	0	0
0.31%	7.24%	7.55%	0	0	0	0	0	0	1

Table 49 : Distribution of collateral types

Calculating "Cumulative Collateral Ratio = Ratio of CollateralType1 + Ratio of CollateralType2" for each region confirms that this finding. Namely Güneydoğu Anadolu Region has the highest Cumulative Collateral Ratio. Then comes Doğu Anadolu and İç Anadolu regions respectively.

This finding is consistent with the finding of Harhoff and Körting (1998) that the bank tend to pledge more collateral requirements of backward regions compared to other counterpart regions.

5.4.5. What if the bank had used its own internal rating for financial distress prediction?

In 2010, the bank has decided to access the credibility of the borrower before setting up the limit and to reevaluate the credit worthiness as long as the borrower is in relation with the bank. The eligibility criteria questionnaire score ranged between 0 and 1000 , and is annually and accumulated in the system for each borrower.

We have ignored internal rating scores of 2010 and 2011 due to the limited number of scores and only used 2012 scores for the prediction. To be able to compare with Model 3 we have sorted the rating scores and starting from the smallest score we have flagged the smallest portion as potential defaulters like in the selectivity ratio of Model 3.

If the bank had used these scores for 1 quarter ahead prediction for financial distress, the results compared to Model 3 would be as follows.

Internal Rating		**Model 3**	
Selectivity Ratio	**Success Rate for Defaulters**	**Selectivity Ratio**	**Success Rate for Defaulters**
22.97	58.33	22.95	72.41
8.48	25.00	8.52	51.72
4.55	8.33	4.58	39.66
3.10	0.00	3.09	25.86
2.07	0.00	2.09	17.24

Table 50 : Performance comparison of bank's internal rating with 1 Quarter ahead prediction of Model 3

With a selectivity ratio of 22.97, internal rating's success for defaulters is %58.33 on the other hand it is %72.41 for Model 3. Similar to different selectivity ratios the success rate for defaulters of Model 3 is always higher than the internal rating results and we may reach a conclusion that Model 3's success rate for defaulters is higher than internal rating outputs for 1 quarter ahead prediction. It is also worth noting that, internal rating system provides no additional value for selectivity ratio smaller than 4, whereas success rate of Model 3 is %25.86 with a selectivity ratio of %3.09.

Internal Rating		Model 3	
Selectivity Ratio	**Success Rate for Defaulters**	**Selectivity Ratio**	**Success Rate for Defaulters**
22.97	46.67	22.95	58.37
8.48	23.33	8.52	40.67
4.55	10.00	4.58	30.62
3.10	3.33	3.09	22.97
2.07	3.33	2.09	17.22

Table 51 : Performance comparison of bank's internal rating with 1 year ahead prediction of Model 3

Similarly, the results of 1 year ahead prediction of internal rating is poorer than Model 3.

6. Conclusion

Developing econometric models is more difficult in social sciences than in natural sciences. This is due to the impossibility of conducting controlled experiments in the former case. This basic fact presents itself under various disguises such as the endogeneity problem, omitted variables bias or simultaneity bias. However when one's main purpose is prediction addressing the issue and solving it only partially _as we did here_ might be acceptable. On the other hand, the use of technology in banking is much more intense compared to other sectors. Thus there is a great deal of information accumulated both in core banking systems and surrounding satellite systems. Utilizing preexisting accounting and loan transactions and other data directly from core banking and satellite systems, and applying the Cox PH method, we developed a financial distress model.

In our study we developed three different models. We tested each one with three different time windows namely [2005, 2010], [2005, 2011] and [2005, 2012]. With each window we performed within sample prediction, one quarter ahead prediction and one year ahead prediction and compared actual defaults with the predicted ones. We used the key performance indicators _*selectivity ratio, success rate for defaulters and calibration ratio*_ to evaluate the best predictive performance of these models. We showed by December 2012, our best performing Model successfully predicts 72.41% of those who default in the first quarter of 2013 with a 22.95% selectivity ratio.

From an applied financial distress prediction perspective this study has three main achievements:

- We have successfully applied Cox's PH – Proportional Hazard approach to financial prediction using real bank loans data and showed a very high degree of predictive accuracy. Note that depending on the bank's operational resources we can adjust the selectivity ratio, which allows the bank to decrease the number of predicted potential defaulters for upcoming time periods with a marginal decrease in predictive accuracy.

- We managed to incorporate macroeconomic variables and firm level time variant information into our model. To the best of our knowledge this is the first such application in the predictive literature applying the Cox methodology to bank loans.

- Compared to existing predictive practices of banks in Turkey, our approach has the following advantage. It is the first application that uses internal bank data pertaining to customer account balances and their change over time. This procedure is both much more accurate and cheaper than existing practice that involves collecting income statements and balance sheets from thousands of customers.

Finally, we believe our approach would provide a higher success rate for predicting defaulters if we had used data from a bank who is the primary bank for its customers. As it happens the small participation bank whose data we were able to access, is not – generally speaking - the primary bank of its customers. As such its transaction accounts are less informative.

7. Appendices

7.1. Model 2

7.1.1. [2005, 2012] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9378193	0.694
Ege Region	1.229825	0.136
Güneydoğu Anadolu Region	0.6658778	0.041
İç Anadolu Region	0.9909208	0.94
Karadeniz Region	0.9895579	0.941
Marmara Region	1.047754	0.664
Fishing Sector	1.490127	0.575
Social Services Sector	0.7213281	0.108
Electricity, Gas and Water Resources Sector	0.3433574	0.005
Real Estate Trading Sector	0.3455676	0.005
Education Sector	0.3015795	0.091
Financial Intermediation Sector	0.2161443	0.126
Mining and Quarrying Sector	0.716763	0.381
Hotels and Restaurants Sector	0.577217	0.101
Defense, Public Administration and Social Security Sector	3.469836	0.08
Health and Social Services Sector	0.4288054	0.001
import	0.6799546	0
export	0.3976653	0.357
dsales_l	0.3911788	0
dsales_m	0.4606269	0
quarterlyaveragebalance	0.9999441	0
numberofcredits	0.8571832	0
numberofdebits	1.002476	0.275
totalrisk	0.9999998	0.006
totallimit	1	0.088
cashriskratio	3.547964	0
totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.001186	0.217
age_at_loan	0.9965826	0
check_length	1.022675	0
numberofbanks	1.090353	0
rtrnedchecksprctn	5.982932	0

cb_returnedchecksprcnt	845535.5	0
w_collateral_1	0.2377678	0.397
w_collateral_2	4.573242	0
w_collateral_3	0.8345116	0.258
tvc_limitutilizationratio	1.061847	0
tvc_numberofstaffquarterly	0.9942338	0

Cox regression -- Breslow method for ties

No. of subjects = 15593 Number of obs = 202615
No. of failures = 1307
Time at risk = 202615
LR chi2(39) = 2673.68
Log likelihood = -10636.506 Prob > chi2 = 0.0000

Note: variables in tvc equation interacted with _t

Table 22 : Cox regression with [2005, 2012] window for Model 2

Event		0	1	Total
0	Count	9,384	4,902	14,286
	Row %	65.69	34.31	100.00
	Column %	97.43	82.23	91.62
	Total %	60.18	31.44	91.62
1	Count	248	1,059	1,307
	Row %	18.97	81.03	100.00
	Column %	2.57	17.77	8.38
	Total %	1.59	6.79	8.38
Total	Count	9,632	5,961	15,593
	Row %	61.77	38.23	100.00
	Column %	100.00	100.00	100.00
	Total %	61.77	38.23	100.00

Table 23 : Actual vs within sample prediction with [2005, 2012] window for Model 2

Actual Defaults in 2013Q1		0	1	Total
0	Count	7,311	3,211	10,522
	Row %	69.48	30.52	100.00
	Column %	99.69	98.92	99.45
	Total %	69.10	30.35	99.45
1	Count	23	35	58
	Row %	39.66	60.34	100.00
	Column %	0.31	1.08	0.55
	Total %	0.22	0.33	0.55
Total	Count	7,334	3,246	10,580
	Row %	69.32	30.68	100.00
	Column %	100.00	100.00	100.00
	Total %	69.32	30.68	100.00

Table 24 : Actual vs 1 quarter ahead prediction with [2005, 2012] window for Model 2

Actual Defaults in 2013		0	1	Total
0	Count	7,242	3,129	10,371
	Row %	69.83	30.17	100.00
	Column %	98.75	96.40	98.02
	Total %	68.45	29.57	98.02
1	Count	92	117	209
	Row %	44.02	55.98	100.00
	Column %	1.25	3.60	1.98
	Total %	0.87	1.11	1.98
Total	Count	7,334	3,246	10,580
	Row %	69.32	30.68	100.00
	Column %	100.00	100.00	100.00
	Total %	69.32	30.68	100.00

Table 25 : Actual vs 1 year ahead predicted with [2005, 2012] window for Model 2

7.1.2. [2005, 2011] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9658313	0.848
Ege Region	1.287625	0.1
Güneydoğu Anadolu Region	0.6620903	0.068
İç Anadolu Region	1.045607	0.737
Karadeniz Region	1.061178	0.701
Marmara Region	1.109508	0.385
Fishing Sector	1.033201	0.974
Social Services Sector	0.5618974	0.019
Electricity, Gas and Water Resources Sector	0.2482727	0.006
Real Estate Trading Sector	0.2573096	0.007
Education Sector	0.3789974	0.171
Financial Intermediation Sector	0.269905	0.191
Mining and Quarrying Sector	0.6686635	0.371
Hotels and Restaurants Sector	0.3942184	0.038
Defense, Public Administration and Social Security Sector	3.588482	0.073
Health and Social Services Sector	0.4701572	0.007
import	0.7006461	0.001
export	0.5320938	0.529
dsales_l	0.3678591	0
dsales_m	0.4858061	0
quarterlyaveragebalance	0.9998462	0
numberofcredits	0.8715247	0
numberofdebits	1.003481	0.177
totalrisk	0.9999997	0.012
totallimit	1	0.016
cashriskratio	3.633489	0
totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.000857	0.551
age_at_loan	0.9962019	0
check_length	1.028098	0
numberofbanks	1.093644	0
rtrnedchecksprctn	5.739352	0
cb_returnedchecksprcnt	491807.3	0
w_collateral_1	0.0046579	0.273
w_collateral_2	4.59321	0
w_collateral_3	0.8478476	0.325
tvc_limitutilizationratio	1.062289	0

tvc_numberofstaffquarterly	0.9959845	0

Cox regression -- Breslow method for ties

No. of subjects = 13620 Number of obs = 161010
No. of failures = 1113
Time at risk = 161010
LR chi2(39) = 2365.99
Log likelihood = -8845.8081 Prob > chi2 = 0.0000
Note: variables in tvc equation interacted with _t

Table 26 : Cox regression with [2005, 2011] window for Model 2

Event		0	1	Total
0	Count	8,782	3,725	12,507
	Row %	70.22	29.78	100.00
	Column %	97.97	80.00	91.83
	Total %	64.48	27.35	91.83
1	Count	182	931	1,113
	Row %	16.35	83.65	100.00
	Column %	2.03	20.00	8.17
	Total %	1.34	6.84	8.17
Total	Count	8,964	4,656	13,620
	Row %	65.81	34.19	100.00
	Column %	100.00	100.00	100.00
	Total %	65.81	34.19	100.00

Table 27 : Actual vs within sample prediction with [2005, 2011] window for Model 2

Actual Defaults in 2012Q1		0	1	Total
0	Count	7,353	2,289	9,642
	Row %	76.26	23.74	100.00
	Column %	99.78	99.13	99.63
	Total %	75.98	23.65	99.63
1	Count	16	20	36
	Row %	44.44	55.56	100.00
	Column %	0.22	0.87	0.37
	Total %	0.17	0.21	0.37
Total	Count	7,369	2,309	9,678
	Row %	76.14	23.86	100.00
	Column %	100.00	100.00	100.00
	Total %	76.14	23.86	100.00

Table 28 : Actual vs 1 quarter ahead prediction with [2005, 2011] window for Model 2

Actual Defaults in 2012		0	1	Total
0	Count	7,272	2,250	9,522
	Row %	76.37	23.63	100.00
	Column %	98.68	97.44	98.39
	Total %	75.14	23.25	98.39
1	Count	97	59	156
	Row %	62.18	37.82	100.00
	Column %	1.32	2.56	1.61
	Total %	1.00	0.61	1.61
Total	Count	7,369	2,309	9,678
	Row %	76.14	23.86	100.00
	Column %	100.00	100.00	100.00
	Total %	76.14	23.86	100.00

Table 29 : Actual vs 1 year ahead predicted with [2005, 2011] window for Model 2

7.1.3. [2005, 2010] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9813858	0.924
Ege Region	1.245556	0.19
Güneydoğu Anadolu Region	0.7400027	0.204
İç Anadolu Region	1.000839	0.995
Karadeniz Region	1.106319	0.541
Marmara Region	1.115652	0.395
Fishing Sector	1.200516	0.856
Social Services Sector	0.5686658	0.031
Electricity, Gas and Water Resources Sector	0.2520687	0.018
Real Estate Trading Sector	0.1702515	0.013
Education Sector	0.4493277	0.26
Financial Intermediation Sector	2.22E-20	.
Mining and Quarrying Sector	0.6752811	0.435
Hotels and Restaurants Sector	0.4054755	0.072
Defense, Public Administration and Social Security Sector	3.889026	0.057
Health and Social Services Sector	0.4121955	0.008
import	0.7119478	0.003
export	0.6319098	0.647
dsales_l	0.368511	0
dsales_m	0.5279672	0

quarterlyaveragebalance	0.9997771	0
numberofcredits	0.8783017	0
numberofdebits	1.001828	0.567
totalrisk	0.9999997	0.034
totallimit	1	0.067
cashriskratio	3.772265	0
totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.000793	0.606
age_at_loan	0.9960769	0
check_length	1.030691	0
numberofbanks	1.103515	0
rtrnedchecksprctn	5.569097	0
cb_returnedchecksprcnt	159336.2	0
w_collateral_1	0.0002213	0.555
w_collateral_2	5.346077	0
w_collateral_3	1.053518	0.767
tvc_limitutilizationratio	1.038845	0.001
tvc_numberofstaffquarterly	0.9972177	0.005

Cox regression -- Breslow method for ties

No. of subjects = 11804 Number of obs = 123758
No. of failures = 965
Time at risk = 123758
LR chi2(38) = 2071.14
Log likelihood = -7485.7808 Prob > chi2 = 0.0000
Note: variables in tvc equation interacted with _t

Table 30 : Cox regression with [2005, 2010] window for Model 2

Event		0	1	Total
0	Count	7,476	3,363	10,839
	Row %	68.97	31.03	100.00
	Column %	98.26	80.15	91.82
	Total %	63.33	28.49	91.82
1	Count	132	833	965
	Row %	13.68	86.32	100.00
	Column %	1.74	19.85	8.18
	Total %	1.12	7.06	8.18
Total	Count	7,608	4,196	11,804
	Row %	64.45	35.55	100.00
	Column %	100.00	100.00	100.00
	Total %	64.45	35.55	100.00

Table 31 : Actual vs within sample prediction with [2005, 2010] window for Model 2

Actual Defaults in 2011Q1		0	1	Total
0	Count	6,580	2,163	8,743
	Row %	75.26	24.74	100.00
	Column %	99.94	99.13	99.74
	Total %	75.06	24.67	99.74
1	Count	4	19	23
	Row %	17.39	82.61	100.00
	Column %	0.06	0.87	0.26
	Total %	0.05	0.22	0.26
Total	Count	6,584	2,182	8,766
	Row %	75.11	24.89	100.00
	Column %	100.00	100.00	100.00
	Total %	75.11	24.89	100.00

Table 32 : Actual vs 1 quarter ahead prediction with [2005, 2010] window for Model 2

Actual Defaults in 2011		0	1	Total
0	Count	6,527	2,120	8,647
	Row %	75.48	24.52	100.00
	Column %	99.13	97.16	98.64
	Total %	74.46	24.18	98.64
1	Count	57	62	119
	Row %	47.90	52.10	100.00
	Column %	0.87	2.84	1.36
	Total %	0.65	0.71	1.36
Total	Count	6,584	2,182	8,766
	Row %	75.11	24.89	100.00
	Column %	100.00	100.00	100.00
	Total %	75.11	24.89	100.00

Table 33 : Actual vs 1 year ahead prediction with [2005, 2010] window for Model 2

7.1.4. Calibration of Model 2

		Risk Score				
		>=1	**>=2**	**>=3**	**>=4**	**>=5**
[2005, 2012] window, Within Sample Prediction	**Selectivity Ratio**	38.23	16.84	9.52	6.45	4.73
	Success Rate	81.03	58.76	42.23	32.90	26.70
	Calibration Ratio	2.12	3.49	4.43	5.10	5.64
[2005, 2011] window, Within Sample Prediction	**Selectivity Ratio**	34.19	15.70	9.37	6.54	4.77
	Success Rate	83.65	62.53	48.88	37.74	30.55
	Calibration Ratio	2.45	3.98	5.22	5.77	6.41
[2005, 2010] window, Within Sample Prediction	**Selectivity Ratio**	35.55	17.53	11.09	7.90	6.01
	Success Rate	86.32	68.39	54.40	44.66	36.48
	Calibration Ratio	2.43	3.90	4.91	5.66	6.07
[2005, 2012] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	30.68	10.78	5.26	2.99	1.97
	Success Rate	60.34	39.66	27.59	15.52	8.62
	Calibration Ratio	1.97	3.68	5.25	5.20	4.38
[2005, 2011] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	23.86	8.30	4.03	2.49	1.48
	Success Rate	55.56	44.44	27.78	13.89	11.11
	Calibration Ratio	2.33	5.36	6.89	5.58	7.52
[2005, 2010] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	24.89	9.88	5.53	3.37	2.28
	Success Rate	82.61	56.52	39.13	13.04	4.35
	Calibration Ratio	3.32	5.72	7.07	3.88	1.91
[2005, 2012] window, 1Year ahead Prediction	**Selectivity Ratio**	30.68	10.78	5.26	2.99	1.97
	Success Rate	55.98	33.49	24.88	14.83	11.48
	Calibration Ratio	1.82	3.11	4.73	4.97	5.84
[2005, 2011] window, 1Year ahead Prediction	**Selectivity Ratio**	23.86	8.30	4.03	2.49	1.48
	Success Rate	37.82	24.36	14.10	8.97	7.69
	Calibration Ratio	1.59	2.94	3.50	3.60	5.21
[2005, 2011] window, 1Year ahead Prediction	**Selectivity Ratio**	24.89	9.88	5.53	3.37	2.28
	Success Rate	52.10	28.57	19.33	5.88	3.36
	Calibration Ratio	2.09	2.89	3.49	1.75	1.47

Table 34 : Calibration of Model 2

7.2. Model 3

7.2.1. [2005, 2012] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9448343	0.728
Ege Region	1.201398	0.186
Güneydoğu Anadolu Region	0.6803541	0.053
İç Anadolu Region	0.9955988	0.971
Karadeniz Region	1.001153	0.994
Marmara Region	1.041789	0.703
Fishing Sector	1.517236	0.558
Social Services Sector	0.7389534	0.136
Electricity, Gas and Water Resources Sector	0.3811825	0.012
Real Estate Trading Sector	0.3672404	0.008
Education Sector	0.2772328	0.07
Financial Intermediation Sector	0.2377047	0.152
Mining and Quarrying Sector	0.7156064	0.379
Hotels and Restaurants Sector	0.5966426	0.124
Defense, Public Administration and Social Security Sector	3.664696	0.068
Health and Social Services Sector	0.4337224	0.001
import	0.659131	0
export	0.4030779	0.365
dsales_l	0.388371	0
dsales_m	0.463915	0
quarterlyaveragebalance	0.999946	0
numberofcredits	0.8589449	0
numberofdebits	1.002523	0.252
totalrisk	0.9999998	0.042
totallimit	1	0.424
cashriskratio	4.041413	0
totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.001151	0.235
age_at_loan	0.9965053	0
check_length	1.022984	0
numberofbanks	1.09247	0
rtrnedchecksprctn	5.832612	0
cb_returnedchecksprcnt	2149286	0
w_collateral_1	0.2701084	0.436
w_collateral_2	5.518333	0

w_collateral_3	1.065594	0.7
limitutilizationratio_precrisis	0.9987387	0.993
limitutilizationratio_crisis	0.7672376	0.167
limitutilizationratio_postcrisis	1.534697	0.005
numberofstaffquarterly_precrisis	0.0000439	0
numberofstaffquarterly_crisis	21981.48	.
numberofstaffquarterly_postcrisis	1.99E-20	.

Cox regression -- Breslow method for ties

No. of subjects =	15593	Number of obs =	202615
No. of failures =	1307		
Time at risk =	202615		
		LR chi2(41) =	2743.16
Log likelihood =	-10601.767	Prob > chi2 =	0.0000

Table 35 : Cox regression with [2005, 2012] window for Model 3

Event		0	1	Total
0	Count	10,047	4,239	14,286
	Row %	70.33	29.67	100.00
	Column %	97.89	79.55	91.62
	Total %	64.43	27.19	91.62
1	Count	217	1,090	1,307
	Row %	16.60	83.40	100.00
	Column %	2.11	20.45	8.38
	Total %	1.39	6.99	8.38
Total	Count	10,264	5,329	15,593
	Row %	65.82	34.18	100.00
	Column %	100.00	100.00	100.00
	Total %	65.82	34.18	100.00

Table 36 : Actual vs within sample prediction with [2005, 2012] window for Model 3

Actual Defaults in 2013Q1		0	1	Total
0	Count	8,136	2,386	10,522
	Row %	77.32	22.68	100.00
	Column %	99.80	98.27	99.45
	Total %	76.90	22.55	99.45
1	Count	16	42	58
	Row %	27.59	72.41	100.00
	Column %	0.20	1.73	0.55
	Total %	0.15	0.40	0.55
Total	Count	8,152	2,428	10,580
	Row %	77.05	22.95	100.00
	Column %	100.00	100.00	100.00
	Total %	77.05	22.95	100.00

Table 37: Actual vs 1 quarter ahead prediction with [2005, 2012] window for Model 3

Actual Defaults in 2013		0	1	Total
0	Count	8,065	2,306	10,371
	Row %	77.76	22.24	100.00
	Column %	98.93	94.98	98.02
	Total %	76.23	21.80	98.02
1	Count	87	122	209
	Row %	41.63	58.37	100.00
	Column %	1.07	5.02	1.98
	Total %	0.82	1.15	1.98
Total	Count	8,152	2,428	10,580
	Row %	77.05	22.95	100.00
	Column %	100.00	100.00	100.00
	Total %	77.05	22.95	100.00

Table 38 : Actual vs 1 year ahead prediction with [2005, 2012] window for Model 3

7.2.2. [2005, 2011] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9768876	0.898
Ege Region	1.247117	0.152
Güneydoğu Anadolu Region	0.6673195	0.073
İç Anadolu Region	1.039927	0.768
Karadeniz Region	1.063787	0.69
Marmara Region	1.090052	0.471
Fishing Sector	1.024618	0.981
Social Services Sector	0.5769006	0.025
Electricity, Gas and Water Resources Sector	0.2874533	0.013
Real Estate Trading Sector	0.2779182	0.011
Education Sector	0.3495571	0.138
Financial Intermediation Sector	0.3179488	0.253
Mining and Quarrying Sector	0.6498401	0.338
Hotels and Restaurants Sector	0.4069279	0.045
Defense, Public Administration and Social Security Sector	3.928149	0.055
Health and Social Services Sector	0.4736993	0.008
import	0.6747297	0
export	0.535325	0.533
dsales_l	0.3622204	0
dsales_m	0.4798541	0
quarterlyaveragebalance	0.9998487	0
numberofcredits	0.8724469	0
numberofdebits	1.003444	0.176
totalrisk	0.9999998	0.091
totallimit	1	0.235
cashriskratio	4.359989	0
totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.000817	0.58
age_at_loan	0.9961153	0
check_length	1.028509	0
numberofbanks	1.096266	0
rtrnedchecksprctn	5.610527	0
cb_returnedchecksprcnt	1209398	0
w_collateral_1	0.0077952	0.319
w_collateral_2	5.884132	0
w_collateral_3	1.205642	0.275
limitutilizationratio_precrisis	0.8502803	0.309

limitutilizationratio_crisis	0.7556198	0.15
limitutilizationratio_postcrisis	1.666971	0.002
numberofstaffquarterly_precrisis	0.0000426	0
numberofstaffquarterly_crisis	22625.82	.
numberofstaffquarterly_postcrisis	1.13E-16	1

Cox regression -- Breslow method for ties

No. of subjects =	13620	Number of obs =	161010
No. of failures =	1113		
Time at risk =	161010		
		LR chi2(42) =	2421.12
Log likelihood =	-8818.2416	Prob > chi2 =	0.0000

Table 39 : Cox regression with [2005, 2011] window for Model 3

Event		0	1	Total
0	Count	9,429	3,078	12,507
	Row %	75.39	24.61	100.00
	Column %	98.30	76.42	91.83
	Total %	69.23	22.60	91.83
1	Count	163	950	1,113
	Row %	14.65	85.35	100.00
	Column %	1.70	23.58	8.17
	Total %	1.20	6.98	8.17
Total	Count	9,592	4,028	13,620
	Row %	70.43	29.57	100.00
	Column %	100.00	100.00	100.00
	Total %	70.43	29.57	100.00

Table 40 : Actual vs within sample prediction with [2005, 2011] window for Model 3

Actual Defaults in 2012Q1		0	1	Total
0	Count	8,099	1,543	9,642
	Row %	84.00	16.00	100.00
	Column %	99.82	98.66	99.63
	Total %	83.68	15.94	99.63
1	Count	15	21	36
	Row %	41.67	58.33	100.00
	Column %	0.18	1.34	0.37
	Total %	0.15	0.22	0.37
Total	Count	8,114	1,564	9,678
	Row %	83.84	16.16	100.00
	Column %	100.00	100.00	100.00
	Total %	83.84	16.16	100.00

Table 41 : Actual vs 1 quarter ahead prediction with [2005, 2011] window for Model 3

Actual Defaults in 2012		0	1	Total
0	Count	8,015	1,507	9,522
	Row %	84.17	15.83	100.00
	Column %	98.78	96.36	98.39
	Total %	82.82	15.57	98.39
1	Count	99	57	156
	Row %	63.46	36.54	100.00
	Column %	1.22	3.64	1.61
	Total %	1.02	0.59	1.61
Total	Count	8,114	1,564	9,678
	Row %	83.84	16.16	100.00
	Column %	100.00	100.00	100.00
	Total %	83.84	16.16	100.00

Table 42 : Actual vs 1 year ahead predicted with [2005, 2011] window for Model 3

7.2.3. [2005, 2010] window

Variable	Hazard Ratio	p-value
Doğu Anadolu Region	0.9763413	0.904
Ege Region	1.206866	0.262
Güneydoğu Anadolu Region	0.7343881	0.193
İç Anadolu Region	0.9946312	0.97
Karadeniz Region	1.086164	0.617
Marmara Region	1.092155	0.493
Fishing Sector	1.21182	0.849
Social Services Sector	0.5784507	0.036
Electricity, Gas and Water Resources Sector	0.284661	0.031
Real Estate Trading Sector	0.1816268	0.016
Education Sector	0.4201908	0.222
Financial Intermediation Sector	2.70E-20	.
Mining and Quarrying Sector	0.6419058	0.379
Hotels and Restaurants Sector	0.409777	0.075
Defense, Public Administration and Social Security Sector	4.328985	0.04
Health and Social Services Sector	0.4126719	0.009
import	0.6951337	0.001
export	0.631186	0.647
dsales_l	0.363052	0
dsales_m	0.5138936	0
quarterlyaveragebalance	0.9997807	0
numberofcredits	0.8799464	0
numberofdebits	1.001632	0.606
totalrisk	0.9999998	0.141
totallimit	1	0.382
cashriskratio	4.529013	0
totallimit_iob	1	0
totalrisk_iob	1	0
limitutilizationratio_iob	1.000749	0.637
age_at_loan	0.9960263	0
check_length	1.029986	0
numberofbanks	1.104922	0
rtrnedchecksprctn	5.451337	0
cb_returnedchecksprcnt	282992.9	0
w_collateral_1	0.0003019	0.565
w_collateral_2	6.630282	0
w_collateral_3	1.500321	0.023
limitutilizationratio_precrisis	0.6857625	0.024

limitutilizationratio_crisis	0.8137339	0.299
limitutilizationratio_postcrisis	1.868964	0.001
numberofstaffquarterly_precrisis	0.000046	0
numberofstaffquarterly_crisis	20910.73	.
numberofstaffquarterly_postcrisis	8.08E-20	.

Cox regression -- Breslow method for ties

No. of subjects =	11804	Number of obs =	123758
No. of failures =	965		
Time at risk =	123758		
		LR chi2(40) =	2139.09
Log likelihood =	-7451.8056	Prob > chi2 =	0.0000

Table 43 : Cox regression with [2005, 2010] window for Model 3

Event		0	1	Total
0	Count	8,049	2,790	10,839
	Row %	74.26	25.74	100.00
	Column %	98.53	76.75	91.82
	Total %	68.19	23.64	91.82
1	Count	120	845	965
	Row %	12.44	87.56	100.00
	Column %	1.47	23.25	8.18
	Total %	1.02	7.16	8.18
Total	Count	8,169	3,635	11,804
	Row %	69.21	30.79	100.00
	Column %	100.00	100.00	100.00
	Total %	69.21	30.79	100.00

Table 44 : Actual vs within sample prediction with [2005, 2010] window for Model 3

Actual Defaults in 2011Q1		0	1	Total
0	Count	7,195	1,548	8,743
	Row %	82.29	17.71	100.00
	Column %	99.94	98.79	99.74
	Total %	82.08	17.66	99.74
1	Count	4	19	23
	Row %	17.39	82.61	100.00
	Column %	0.06	1.21	0.26
	Total %	0.05	0.22	0.26
Total	Count	7,199	1,567	8,766
	Row %	82.12	17.88	100.00
	Column %	100.00	100.00	100.00
	Total %	82.12	17.88	100.00

Table 45 : Actual vs 1 quarter ahead prediction with [2005, 2010] window for Model 3

Actual Defaults in 2011		0	1	Total
0	Count	7,139	1,508	8,647
	Row %	82.56	17.44	100.00
	Column %	99.17	96.23	98.64
	Total %	81.44	17.20	98.64
1	Count	60	59	119
	Row %	50.42	49.58	100.00
	Column %	0.83	3.77	1.36
	Total %	0.68	0.67	1.36
Total	Count	7,199	1,567	8,766
	Row %	82.12	17.88	100.00
	Column %	100.00	100.00	100.00
	Total %	82.12	17.88	100.00

Table 46 : Actual vs 1 year ahead predicted with [2005, 2010] window for Model 3

7.2.4. Calibration of Model 3

		Risk Score				
		>=1	**>=2**	**>=3**	**>=4**	**>=5**
[2005, 2012] window, Within Sample Prediction	**Selectivity Ratio**	34.18	16.49	10.41	7.52	5.80
	Success Rate	83.40	63.58	49.43	39.63	32.82
	Calibration Ratio	2.44	3.86	4.75	5.27	5.66
[2005, 2011] window, Within Sample Prediction	**Selectivity Ratio**	29.57	15.39	10.15	7.35	5.81
	Success Rate	85.35	67.65	54.45	44.03	37.38
	Calibration Ratio	2.89	4.40	5.37	5.99	6.43
[2005, 2010] window, Within Sample Prediction	**Selectivity Ratio**	30.79	16.58	11.13	8.42	6.66
	Success Rate	87.56	70.88	56.79	47.15	40.41
	Calibration Ratio	2.84	4.28	5.10	5.60	6.07
[2005, 2012] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	22.95	8.52	4.58	3.09	2.09
	Success Rate	72.41	51.72	39.66	25.86	17.24
	Calibration Ratio	3.16	6.07	8.65	8.37	8.25
[2005, 2011] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	16.16	6.21	3.68	2.36	1.64
	Success Rate	58.33	52.78	38.89	25.00	13.89
	Calibration Ratio	3.61	8.50	10.57	10.61	8.45
[2005, 2010] window, 1 Quarter ahead Prediction	**Selectivity Ratio**	17.88	7.63	5.00	3.58	2.56
	Success Rate	82.61	60.87	56.52	34.78	21.74
	Calibration Ratio	4.62	7.98	11.31	9.71	8.51
[2005, 2012] window, 1Year ahead Prediction	**Selectivity Ratio**	22.95	8.52	4.58	3.09	2.09
	Success Rate	58.37	40.67	30.62	22.97	17.22
	Calibration Ratio	2.54	4.78	6.68	7.43	8.25
[2005, 2011] window, 1Year ahead Prediction	**Selectivity Ratio**	16.16	6.21	3.68	2.36	1.64
	Success Rate	36.54	25.64	18.59	12.18	8.33
	Calibration Ratio	2.26	4.13	5.05	5.17	5.07
[2005, 2010] window, 1Year ahead Prediction	**Selectivity Ratio**	17.88	7.63	5.00	3.58	2.56
	Success Rate	49.58	29.41	24.37	16.81	9.24
	Calibration Ratio	2.77	3.85	4.88	4.69	3.62

Table 47 : Calibration of Model 3

8. References

Abdel-Khalik, A. R. (1993). Discussion of "Financial Ratios and Corporate Endurance: A Case of the Oil and Gas Industry". *Contemporary Accounting Research, Vol. 9 Issue 2*, p. 695-705.

Ahmad, A. (1993). Contemporary Practices of Islamic Financing Techniques. *Islamic Development Bank, Islamic Research and Training Institute*.

Ang, J., Lin, J. W., & Tyler, F. (1995). Evidence on the Lack of Separation between Business and Personal Risks among Small Businesses. *The Journal of Entrepreneurial Finance, 4*(2), p. 197-210.

Avery, R., Bostic, R., & Samolyk, K. (1998). The role of personal wealth in small business finance. *Journal of Banking & Finance*(22), p. 1019-1061.

Aysan, A. F., Dolgun, M. H., & Turhan, M. (2014). Assessment of the Participation Banks and Their Role in Financial Inclusion in Turkey. *Emerging Markets Finance and Trade*, p. 99-111.

Brynjolfsson , E., & Saunders, A. (2013). *Wired for Innovation: How Information Technology Is Reshaping the Economy.*

Casu, B., Girardone, C., & Molyneux, P. (2006). *Introduction to Banking.* Prentice Hall.

Cevik, S., & Charap, J. (2011). The Behavior of Conventional and Islamic Bank Deposit Returns in Malaysia and Turkey.

Chong, B. S., & Liu, M.-H. (2009). Islamic banking: Interest-free or interest-based? *Pacific-Basin Finance Journal*, p. 125-144.

Çokgezen, M., & Kuran, T. (2015). Between Consumer Demand and Islamic Law : The Evolution of Islamic Credit Cards in Turkey. *Journal of Comparative Economics*, p. 862-883.

Davutyan, N. (2008). Estimating the size of Turkey's informal sector: an expenditure-based approach. *Journal of Economic Policy Reform, 11*(4), p. 261–271.

Davutyan, N., & Öztürkkal, B. (2015). Determinants of Saving-Borrowing Decisions and Financial Inclusion in a High Middle Income Country: the Turkish Case.

Dedrick, J., Gurbaxani, V., & Kraemer, K. (2003). Information Technology and Economic Performance: A Critical Review of the Empirical Evidence. *ACM Computing Surveys, 35*(1), p. 1-28.

Golin, J., & Delhaise, P. (2013). *The Bank Credit Anaysis Handbook.* Wiley.

Harhoff, D., & Körting, T. (1998). Lending relationships in Germany - Empirical evidence from survey data. *Journal of Banking & Finance*(22), p. 1317 - 1353.

Heffernan, Shelagh. (2005). *Modern Banking.* Chichester: John Wiley & Sons Ltd.

Helwege, J. (1996). Determinants of Savings and Loan Failures: Estimates of a Time-varying Proportional Hazard Function. *Journal of Financial Services Research, 10 4*, p. 373-392.

http://www.tkbb.org.tr/documents/TKBB_Strateji_Belgesi_Ingilizce.pdf. (2015). Retrieved December 2015, from Participation Banks Association of Turkey.

Jappelli, T., & Pagano, M. (2002). Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking & Finance*, p. 2017–2045.

Kaya, R. (2013). *Türk Bankacılık Sisteminde Katılım Bankalarının Tercih Nedenleri ve Bir Araştırma.* Istanbul: Marmara Üniversitesi : Bankacılık ve Sigortacılık Enstitüsü Bankacılık Anabilim Dalı.

Kewin C.W. Chen, C.-W. J. (1993). Financial Ratios and Corporate Endurance: A Case of the Oil and Gas Industry. *Contemporary Accounting Research, Volume 9*(Issue 2), p. 667-694.

Khan, F. (2010). How 'Islamic' is Islamic Banking? *Journal of Economic Behavior & Organization*, p. 805-820.

Kleinbaum, D. G. (1996). *Survival Anaysis A Self - Learning Text.* Springer Science.

Kuran, T. (2004). Islam and Mammon : The Economic Predicaments of Islamism.

Leclere, M. (2008). *Survival Analysis and Omitted Dividents, in Advances in Credit Risk Modelling and Corporate Bankruptcy Prediction.* (S. Jones, & D. Hensher, Eds.) Cambridge University Press.

Li-Sheng Chen, M.-F. Y.-M.-S.-M.-S.-H. (2005). Predictive Survival Model with Time-dependent Prognostic Factors: Development of computer-aided SAS Macro Program. *Journal of Evaluation in Clinical Practice, Vol. 11 Issue 2*, p. 181-193.

Mester, L., Nakamura, L., & Renault, M. (2007). Transactions Accounts and Loan Monitoring. *The Society for Financial Studies, 20*(3), p. 529-556.

Mirakhor, A., & Krichene, N. (2009). The Recent Crisis: Lessons for Islamic Finance. *IFSB 2nd Public Lecture on Financial Policy and Stability*.

Nam, C. W., Kim, T. S., Park, N. J., & Lee, H. K. (2008). Bankruptcy Prediction Using a Discrete-time Duration Model Incorporating Temporal and Macroeconomic Dependencies. *Journal of Forecasting, Vol. 27 Issue 6*, p. 493-506.

National Automobile Dealers Association, Monthly Sales Trends. (2009). *AutoExec*, p. 24.

Norden, L., & Weber, M. (2010). Credit Line Usage, Checking Account Activity, and Default Risk of Bank Borrowers. *The Society for Financial Studies*, *23*(10), p. 3665-3699.

Oh, S. A., Nam, C. W., Kim, T. S., & Lee, H. K. (2013). Assessing the Proportionality Assumption in Deafult Rate Forecasting Using the Proportional Hazard. *KAIST Business School*.

Oliner , S., & Sichel, D. (2002). Information Technology and Productivity: Where Are We Now and Where Are We Going? Retrieved from http://www.federalreserve.gov/pubs/feds/2002/200229/200229pap.pdf

Ongena, S., & Sendeniz Yuncu, İ. (2011). Which firms engage small, foreign, or state banks? And who goes Islamic? *Journal of Banking & Finance*.

Özsoy, İ., Görmez, B., & Mekik, S. (2013). Reasons for Preference of Participatory Banks in Turkey: An Empirical Investigation. *Yönetim ve Ekonomi , 20*(1).

Robert J. Kauffman, B. W. (2008). Tuning into the Digital Channel: Evaluating Business Model Characteristics for Internet Firm Survival. *Information Technology & Management, Vol.9 Issue 3*, p. 215-232.

Rose, P. S., & Hudgins, S. C. (2008). *Bank Management and Financial Services* (7. ed.). New York: The McGraw-Hill Companies.

Sarac, M., & Zeren, F. (2015). The dependency of Islamic bank rates on conventional bank interest rates: further evidence from Turkey. *Applied Economics, 47*(7), p. 669–679.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business, Vol. 74 Issue 1*, p. 101.

Singer, J. D., & Willett, J. B. (2003). *Applied Longittudinal Data Analysis.* Oxford University Press.

Whalen, G. (1991). A proportional hazards model of bank failure: An examination of its usefulness as an early warning. *Economic Review*, Vol. 27, Issue 1.

Wheelock, D. C. (1995). Explaning Bank Failures: Deposit Insurance, Regulation, and Efficiency. *Review of Economics & Statistics, Vol. 77 Issue 4*, p. 689.

Williamson, S., & Johnston , L. (2015). *What was the U.S GDP Then?* Retrieved December 2015, from http://www.measuringworth.org/usgdp/