Fall 12-18-2015

An Applied Credit Scoring Model and Christian Mutual Funds Performance

Esther E. Castro
University of New Orleans, esther15castro@gmail.com

An Applied Credit Scoring Model and Christian Mutual Funds Performance

A Dissertation

Submitted to the Graduate Faculty of the
University of New Orleans
In partial fulfillment of the
Requirements for the degree of

Doctor of Philosophy
In
Financial Economics

by

Esther Castro

B.S. LeTourneau University, 2009
M.S. University of New Orleans, 2013

December, 2015

Dedication

I would like to dedicate my dissertation first and foremost to my God for whom I live and breathe. Without His love and grace, none of this would have been possible. He has sustained me throughout this journey and continue to shower me with blessings. Thank You Lord Jesus!

Also to my parents, Juan and Lizete, who has love and supported me my whole life. They have are a living example of what God's grace and hard work can accomplish. Thank you for teaching me to strive for excellence.

Acknowledgement

My Ph.D. journey would not have been possible without the prayer and support of friends and family in my life. Thank you to everyone who has prayed for me throughout my process and while I was defending, I am sincerely humbled by your love. I also would like to thank my dissertation chair, Dr. Mohammad K. Hassan for guiding me through my dissertation process. Your support and advice has been invaluable and I could not have asked for a better chair. Thank you Dr. Jose Francisco Rubio for your help and guidance in my essay. A special thank you to Dr. Rosa and Jefferson Financial Credit Union for which data was provided. Without their assistance this project would not have been possible. Last, but not least, a special acknowledgement to the UNO Faculty, both in the Financial Economics Department as well as the Management and Marketing department for whom I have had the pleasure to work with.

Table of Contents

List of Figures

List of Tables

Abstract

This dissertation comprises two different financial essays. Essay 1, “An Applied Credit Score Model,” uses data from local credit union to predict the probability of default. Due to recent financial crisis regulation has been enacted that makes it essential to develop a probability of default model that will mitigate charge-off losses. Using discriminant analysis and logistic regression this paper will attempt to see how well credit score can predict probability of default. While credit score does an adequate job at classifying loans, misclassification of loans can be costly. Thus while credit score is a predictor, there is danger in relying solely on its information. Thus other variables are needed in order to more accurately be able to find the probability of default. Essay 2, “Christian Mutual Fund Performance,” draws attention to a much ignored type of funds, Christian mutual funds. The following questions are asked: How does Christian mutual fund perform compared to the market? Is there a difference in performance during recessions as indicated by literature? Is Christian mutual fund performance different than SRI funds? How do Catholic and Protestant fund perform? Looking at qualitative evidence, Christian mutual funds place much more importance on moral issue than SRI funds. Thus there is a clear difference in objectives and the type of screening that these two mutual fund pursue. Overall data reflects that screened data perform worse than the market, however during recession screened funds perform as well and at times better than the market. Christian mutual funds tends to perform worse than SRI funds.

Keywords: Credit Score, Probability of Default, Loans, Mutual Fund Performance, Religion

Chapter 1

Essay 1: An Applied Credit Scoring Model

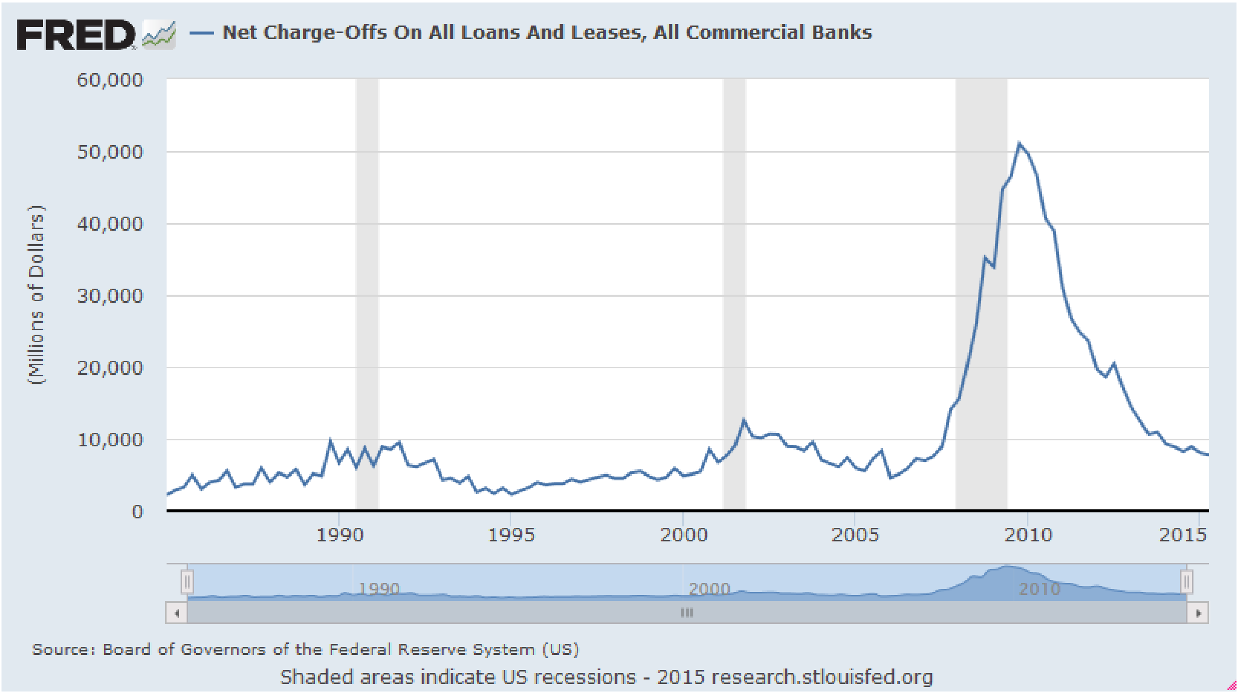
1. Introduction

The banking industry has gone through several changes in the last 60 years. These changes have in part to do with regulatory changes and financial product innovation. Yet one thing has remained: the demand and dominance of consumer lending. Consumer credit loans have increased in the banking industry, in general, as well as in Credit Unions in the last 60 years. Consumer loans have contributed to the way of life for many Americans. For many Americans who have wanted to increase their standard of living, consumer loans have been the answer. Research has shown that consumer loan is among the most profitable loan a bank can make. However, Functional Cost Analysis (FCA) program conducted by the Federal Reserve found that consumer loans are among the most risky and costly loanable funds that a bank grants to their customer. Recovering a loan is dependent upon the consumer's economic state, heath state, and many times moral character. Consumer loans, furthermore, are ascertained to be cyclical with the overall state of the economy. With this uncertainty surrounding consumer lending, it poses a challenge for banks to predict loan portfolio risk. The recent subprime crises accentuate the need for measuring the portfolio risk of banks. Capturing the risk for their mortgages, small business loans, or individual borrowers influences the financial institution in making appropriate interest rate, lending policy, and reserve requirement changes.

There are different types of consumer loans: residential mortgage, non-residential loans, and credit card loans. Consumer loans are a very profitable form of loans since they are usually priced well above the cost of funding them. However, since the financial state of the consumer can fluctuate due to illness or loss of employment, consumer credit is also among the most risky and costly product for banks. For this reason, interest rates are set high for most consumer loans. Banks thus must be prepared for the event that loans may not be collected. Charge-off can be defined as an amount of debt that is unlikely to be recovered, thus must be written-off. According to the Federal Reserve as seen in figure 1, commercial banks' net charge-off reached a peak of almost $51 billion in the fourth quarter of 2009, since then it has declined so that in the first quarter of 2015 the net charge-off is around $8 billion. This, however, is still a considerable sum of money. Therefore it is critical to be able to predict the possibility of charge-off and the likelihood of default.

<u>Figure 1: Net Charge-Off</u>

The Net Charge-offs on all loans and leases for all commercial banks reported quarterly and not seasonally adjusted. The figures are in millions of dollars.



One major advantage that smaller community banks and credit unions have over large banks is their relationship with customers. A key factor when analyzing a consumer's loan application is to have knowledge of the borrower's character and ability to pay. Knowing the person's character and thus their sense of moral responsibility is a good indication of their intentions to pay back. A consumer loan officer also should seek insight of customer's credit history. There are over 2,000 credit bureaus in the United States that provide credit rating for most individuals who has at one time or another, borrowed money. Many banks use credit scoring system to evaluate their loan application. This system has a major advantage that it can sift through large quantities of credit application with minimal labor, thus reducing operating costs. A bank establishes a cutoff point which would yield the greatest net savings in loan losses. Yet credit scores provide limited information and should not be the only precursor used. Due to their small size, small banks do not have the same resources as larger banks to calculate portfolio loan loss.

The current crisis of the subprime mortgages emphasizes the need to have simple models that are capable of capturing the financial institution population risk. The idea is to find and pre-identify certain factors that determine the probability of default for a given loan or credit by using quantitative scores. In some cases, the score can be interpreted as a probability of default. The score may be used to classify or quantify the potential of default or to group a borrower into a "good" or "bad" category. Credit scoring systems are also known as behavioral scoring, in which scores try to predict behavioral trends exhibited by customers. Credit scoring applies logic to behavioral results and provides warning reports to portfolio management personnel on credits that possess undesirable behavioral attributes deemed to be associated with greater potential loss. Attributes of credit scoring systems may include, but are not limited to, updates of loan accounting system information, updates of loan deposit information, and updates of information from personal and/or business credit bureau files. With a credit scoring system, accounts can be queued to portfolio management personnel for risk grade establishment and exposure assessment.

The purpose of this study is to aid small community banks and credit unions in constructing a model that will predict portfolio loan risk. This study will define consumer lending and highlight some interesting statistics relating to the current status of consumer lending in the banking industry. We will also analyze different types of models and methodology that has been used in the past. This study will provide valuable information to the portfolio manager of a bank which is essential to making the correct decision regarding consumer loans. We found that credit score is able to accurately identify default loans by 85%. However, they misclassify loans that defaulted as pay-off loans by almost 15%. Thus, using credit score as a sole predictor of default can be costly. Credit score can only explain 43% of probability of default. And thus other variables need to be included.

2. Literature Review

Tufano (2009) defines consumer finance as "the study of how institutions provide goods and services to satisfy the financial functions of households, how consumers make financial decisions, and how government action affects the provision of financial services." Although in academic research corporate finance overshadows consumer finance in asset value the consumer sector dominates the corporate sector. The recent economic crisis attests to the importance of the consumer sector. Though there may be several different factors that contributed to the economic recession, without a doubt the subprime mortgage market played a big role. Thus it is important to understand consumer finance. Banks play a large role in consumer lending yet one of their greatest challenges is to find ways to prevent loan losses. Therefore a probability of default model for portfolio loan is essential.

Risk Management

Regulations and risk management procedures are another important reason for probability of default models. One key component that bank managers are concerned with is asset management. Managers are tasked to minimize risk by diversifying their portfolio and acquiring assets with low default risk. A big part of this is managing credit risk.

One standard that has been put emplace to address credit risk was formed by the G-10 central banks that established the Basel Accords. Basel Accords establish uniform capital requirements across nations in order to "strengthen the soundness and stability of the international banking system" and to decrease "competitive inequality among international banks" (Basle, 1988). Basel I agreement was issued in 1988 and focuses on capital requirements on assets with differing credit risks. Basel I set up a framework on how to categorize asset into five risk categories (0, 10, 20, 50 and 100%). An international operating bank is required to have 8% or less risk weight. Basel II, a revision of the previous framework, was agreed upon in June 2004, but was going to be implemented in late 2007; however the financial crisis interrupted full implementation. The main objective of Basel II was to revise the previous framework to be more risk-sensitive (Basle, 2006). They developed a three pillar concept: 1. Minimum capital requirements, 2. Supervisory review, and 3. Market discipline. While the previous provision focused on credit risk, this new agreement also ensures that operational risk and market risk be quantified along with credit risk in order to have the appropriate capital adequacy in banks. Supervisor develops review procedures that use assessment of risk tools in order to ensure that banks have adequate capital. Disclosure is also an important factor of the third pillar. Due to the disruption of the implementation from the financial crisis, a new framework was drafted, but as of yet not fully implemented. The new accord's objective is to make banks more resilient and able to absorb financial and economic stress shocks (BIS 2011). This new accord thus attempts to strengthen the three pillars by improving risk management, governance and disclosure. The committee also reinforces fundamental microprudential regulations as well as introduced macropruedential regulations.

After the Financial Crisis, stress testing programs have been put in place in order to test bank's ability to react to stressful situations, such as an economic crisis. The main objective is to put banks in a hypothetical hostile condition in order to ascertain up to what point the bank will be able to remain afloat. Regulatory agencies, as well as individual banks, may perform these tests in order to determine their weak spots in order to take corrective action. While stress test look at ten different factors, one factor is the bank's exposure to default. Banks may use their own risk management default model in order to find their exposure at defaults. Two primary factors that are used is the probability of default (the likelihood that the borrower will not be able to pay back the loan) and the loss given default (the loss that a bank endures due to borrower default on the loan). Thus it is becoming increasingly important that a bank create their own model to predict default.

Classical Probability of Default Models

Since bank loan information is hard to come by, probability of default is generally modeled using corporate securities, specifically bonds. There has been a different progression of probability of default models over the years. There are two principal models in literature regarding corporate default: structural approach pioneered by Merton (1974) and reduced-form approach developed by Jarrow & Turnbull (1995). Merton (1974) was one of the first to develop a model of probability of default for bonds. His purpose was to develop a theory in which he could price bonds that had significant probability of default. Merton's model was of the first generation that linked the probability of default with firm's asset volatility and leverage. Thus the probability of default is driven by the company's asset value and its variability. However, Merton assumes that a default can only occur at maturity, has limitation in bond contract, and it dismisses the possibility of reorganization. Merton's framework was adopted soon after by many researchers such as Geske (1977), Smith & Warner (1979), and Black & Cox (1976). Geske extended the model to show that risky securities can be valued as compound options. Smith & Warner also employs the model to investigate the relationship between the ways debt contracts are written and the conflict between stockholders and bondholders. Black & Cox sought to expand Merton's model by fixing some of Merton's simplified assumption. In their extended model they explored the effect of three different, yet standard, bond indentures: safety covenants, subordination arrangements, and restrictions on the financing of interest and dividend payments. They found that these provisions positively affect the value of the bond. Their model also takes into account bankruptcy cost. Both Merton's and Black & Cox's extension however still assumes that interest rate is constant. Longstaff & Schwartz (1995) strived to remedy this by incorporating both default risk and interest rate risk to both closed-form valuation of floating rate and fixed rate debt. Even with all these improvements, the structural model has one huge drawback according to Duffie & Singleton (2003) the firm's assets are neither traded nor observed. Jarrow & Turnball (1992, 1995) thus developed an alternative model which sought to address this problem. The reduced form model, as it later became known, values stock using a stochastic process that takes exogenously both the default-free term structure and the risky debt term structure. Jarrow, Lando, & Turnbull (1997) extend the model by presenting a finite state using Markov chain model in the firm's credit rating. Lando (1998) generalizes Jarrow et al. model using Cox process. This framework allows dependency between credit risk and market risk factors. This model thus reduces the technical difficulties caused by default correlations. While both models are helpful, they both have drawbacks. Zhou (2001) thus attempts to combine the models in such a way that it would retain the advantages of both the structural and reduced-form approach. The structural model followed a diffusion process which does not allow a sudden drop in firm value; the reduced-form approach however regards default as only an unpredictable Poisson event. Thus Zhou's framework includes both default risk and interest rate risk and allows for the default to possess both a continuous and a jump component. Jarrow (2009) writes a comprehensive paper comparing the structural and the reduced form models and concludes that the reduced form model is the better credit risk model. Another area of research that is related to calculating the probability of default is to valuating the recovery rate in the event that default occurs.

Banks

While much work has been done on measuring the probability of default and recovery rate on bonds and options, much less have been done on bank loans. Yet bonds and bank loans are monitored differently. While loans are monitored by bank managers, bonds must be monitored by the public who holds them. There is a dichotomy of information between the two. Where the bank has superior resources and information on their borrower, bondholder usually does not. Thus there is a monitoring advantage of banks over bondholders (Diamond 1984). Recently, there has been a growth in loan securitization. This has brought about a fear of monitoring deterioration and moral hazard behavior. Loan securitization could give the bank the opportunity of getting rid of "lemon" loans and keep only the best quality loans. However, moral hazard can be mitigated, and banks still have a monitoring advantage compared to public bonds (Altman, Gande, Saunders, 2010). In fact Altman et al. found evidence that loan returns Granger cause bond returns before firm defaults on its loans. Altman & Suggitt (2000) assess the default rate experience on large, syndicated bank loans. According to them, the most fundamental aspect of credit risk models is the rating of the underlying credit asset and the associated expected and unexpected risk migration patterns. The mortality rates on bank loans are extremely similar to corporate bonds, but loan default rates appear to be noticeably higher than bonds for the first two years after issuance. Thus, in the first two years after issuance, loan default rates are higher than bond default rates.

A very interesting result is found using loans from a Spanish credit institution from 1988-2000 (Jimenez, & Saurina, 2004). They analyze the determinant of probability of default focusing on three variables: collateral, type of lender, and bank-borrower relationship. They find that collateral for a loan actually increases the probability of default of a loan. While this may sound counterintuitive, their theory relates that banks tend to screen less on a loan in which collateral is provided. The risk of default was found to be affected by the type of lender, or the type of bank giving the loan. The model revealed that savings banks have a higher risk compared to commercial banks' loans. One explanation for this is that the savings' banks are controlled by managers, as opposed to commercial banks which are controlled by shareholders. Regarding relationship banking, it was found that the closer the relationship between the firm and the bank, the higher the risk of default. If a firm is being financed by only one bank and thus shows a greater commitment to that bank, the bank will be more likely to take on the risk, and thus the probability of default is higher. This paper thus strongly encourages the use of thorough screening process when making loan decisions in order to avoid high default rates.

Jacobson, & Roszbach (2003) create a model that determines a bank's decision on whether or not to approve a loan and the borrower's risk of default on that loan. The researchers used loan applications collected from a Swedish lending agency between September 1994 and August 1995. The data included the number of the applicant, the date submitted, size of the loan, status on whether the loan was good or bad as of October 1996, and what date the loan reached the bad status also other demographical information was included such as sex, marital status, residence, citizenship, age, income, wealth, and homeownership status. After discarding variables due to endogeneity issues, they found that the income level of the applicant, whether the applicant owns a house, whether the applicant has taxable income from a business, loans outstanding, and the existence of a guarantor all have a positive effect on whether or not the applicant gets approved for a loan. The income variable stood out because even though an applicant with a higher income was more likely to have his or her loan approved, the applicant was more likely to default on the loan. Several of the variables that affected the loan approval decision do not affect the loan's risk of default. Another interesting find was that the size of the loan itself does not alter the loan's chance of being defaulted on. A borrower was no more likely to default on a larger loan than default on a small loan. Looking at portfolio loans for a Farm Credit District, three ratios was identified to be significant in explaining the probability of default for the data set loans: repayment capacity, owner equity, and working capital (Featherstone & Roessler, & Barry, 2006). Another discovery that was made was that as loans age increase, the probability of default decreases. This is a logical result since the longer a loan continues without defaulting, the more stable the payments have been for a longer period of time, and the default rate will be lower.

Featherstone, Roessler, and Barry (2006) conducted research on credit analysis. The primary risk that financial institutions face is credit risk, and thus they most perform some-type of risk-taking systems. These ratings serve multiple purposes, including contributing to the loan origination process, aiding in monitoring the safety and soundness of loan portfolios, and in management reporting, facilitating adequacy of loan reserves, and providing components of loan pricing profitability analysis systems. The “fundamental goal” of a credit risk-rating system is to accurately estimate the credit risk of a specific transaction or portfolio of transactions/assets. The “ultimate goal” is to measure the expected and unexpected loss from investing in an asset and the capital required to support it. Default mode and mark-to-market are the two main approaches to measuring credit risks identified in the literature (Barry 2001). The default mode approach “focuses directly on the possibilities of loan loss, including probability of default and the severity of loss given that default has occurred.” The mark-to-market approach attempts to measure how future changes in the credit risk characteristics of a loan or a group of loans will affect the loan(s) market value, including potential losses in value. Furthermore, the new Basel Accord may allow lenders to benefit from a more accurate risk rating of their loan portfolio. The goals of the Basel Accord are to tailor risk management of the financial institution and to increase segmentation of the loan portfolio by risk rating. Wilson (2000) found that banks are more likely to fail if they have low capitalization, higher ratios of loan to asset, and poor quality of loan portfolios. Lopez (2002) looks at the relationship between firm’s probability of default, asset size and average asset correlation. He found that the average asset correlation negatively correlates with the probability of default. Thus the probability of default is not closely related with the macroeconomic environment instead default is mainly due to idiosyncratic factors. This brings into question whether Basel accords regard for other type of risks are warranted. Wheelock and

The Great Recession

The importance of consumer lending became obvious during the Great Recession. One key factor of the Great Recession was due to the increase subprime and near-prime lending, which was further aggravated by the securitization of these loans. The Financial Crisis of 2008 followed similar trends to other crisis (Demyanyk & Van Hemert, 2011). First, there was an evident boom in the subprime mortgage market. Second, a bust occurred in 2007 which is signaled by house foreclosures, high delinquencies and default rates. The subprime crisis led to spill over into other credit markets. The crisis intensified when underwriting standards deteriorated along with loan quality which led to an increase in loan risk that was not reflected by an increase in price, which led to a collapse in the market.

Kwan (2001), using data from nine years from Merrill Lynch and Fannie Mae , finds the average annual growth rate of subprime mortgages was 26 percent. Kwan concludes that subprime loans can affect credit values and the loans that are tied in with them. With an increase in subprime lending in the 2000s due to predatory practices, it was only a matter of time for a banking crisis to occur. The credit boom emanated from 2001 – 2006 and bust in 2007, mainly due to the large subprime securitized mortgage market (Demyanyk & Hasan, 2010). In 2008, the subprime securitized mortgage market was roughly around $1.8 trillion which is about one-third of the securitized market and 16% of total mortgage debt. Though many people doubted that such a comparatively small market could induce such a crisis, the complexity, however, of the innovated security contributed to the collapse. Keys et al. (2008) studied the link between securitization and screening subprime mortgage backed securities. They found that lenders that are most likely to securitize portfolios have less motivation to screen borrowers and more likely to default (by 10-25%) compared to those portfolios with similar risk but with less probability of securitization. Furthermore, Mian & Sufi (2008) revealed a positive relationship between securitization and subprime lending and subsequent defaults. In geographical zones where borrowers were once denied credit (in 1996) received an exceptional growth of accepted credit and later mortgage default. In congruent with the growth in mortgage credit in this area and decrease in income growth, there was an increase to securitize these subprime mortgages. In 2008, Bernanke informed the public that 10% of near-prime mortgages and over 20% of subprime mortgages were delinquent and 2.25 million foreclosures were initiated. In 2009, these figures increased to 13% and 25% respectively. While many maligned the nontraditional features involved in mortgage contracts, Mayer, Pence and Sherlund (2009), found that the biggest reason delinquency rates were remarkably unmanageable was because it was originated to borrowers with low credit score and high loan-to-value ratios. LaCour-Little & Zhang (2014) looked at estimating the probability of default and loss given default for home equity loans around the time of the financial crisis. In this paper, they compiled data from large commercial banks, where loans were originated during 2004-2008 and tracked from 2008-2012. They are particularly interested in the relationship between loan outcomes and the lender decision to securitize the asset. After they examined loan performances, including LGD for home equity loans they ascertained that there was an increase in the probability of default among the particular loans that were securitized. Lending to the corporate sector through loan syndication also suffered during the 2008 Financial Crisis (Ivashina & Scharfstein, 2010). There was a 37% drop in lending during September through November period prior to the past three month period and 68% decline since the peak in 2007. The authors expostulate however, that a decrease in lending does not necessary mean a reduction in credit supply. A decrease in lending is due to a reflection of the increase in risk. However, they noted that banks with a “strong base of deposits” will cut their lending less. For example, in August-November 2008 period, the median range bank reduced lending by 38% while a bank with a deposit of one standard deviation below (above) reduced their lending by 51% (14%). Thus a bank with a solid deposit intake are inherently less risky and are capable of lending even through the crisis.

Credit Scores Literature

Before the emergence of credit score, credit worthiness was measured in various ways, but normally boiled down to a judgment call (Fensterstock, 2003). A loan officer would base their decision off a system that captures the borrowers Character, Capacity, Capital, and Condition; also known as the four C’s. Saunders & Allan (2010) includes another C, Collateral. Other than these factors, managers also had to take interest rate into account. Loan managers are aware of the nonlinear relationship between interest rate and expected return on loan. If interest rates are at a relatively “low” level, by increasing the rate, return should also increase. However if interest rates are relatively high, the expected return on loan decreases due to adverse selection and risk shifting. However, due to the subjective nature of loan decisions, individual or business credit worthiness could vary drastically depending on the loan officer. The judgmental system uses internal and external credit experience within a formula to determine the score. This method looks at the customer’s payment history, credit agency ratings, and financial statements among other factors. However, it is inefficient because it takes copiousness amount of manual work, especially in the initial set-up of the system. Also, the weights used can be biased because of irrelevant factors that should not weigh into the situation. It is difficult of the judgmental system to determine where errors are originating from, which makes it difficult to update and correct the system.

Now, individual's risk assessment is usually given by their credit score, which is calculated by credit bureaus, Fair Isaac Corporation (FICO) being the most common. These credit scores are developed using predictive algorithms that use personal information to estimate an individuals' risk (Citron & Pasquale, 2014). FICO was created in 1956, and developed a three-digit credit score system which scores ranged from 300 to 850 where the lower the score, the more likelihood the individual would default. In many instances, credit scores are used to price loans in order to remain objective. According to FICO, their scores are calculated using credit data which are grouped into five different categories: 1) Payment history; 2) Amounts owed; 3) Length of credit history; 4) New credit; and 5) Types of credit used.

Credit scores are calculated by determining which factors are pertinent to the score and multiplying them by a respective weight of importance (Fensterstock, 2003). Credit scores can get much more complex than that, however, there are four main kinds of credit scoring systems, including the judgmental system (which we previously mentioned), the neural network-based system, the statistical-based system, and the genetic algorithm-based system. The last three are scientific-based and can be up to 30% more accurate than the human judgment system; they also meet requirements set by Sarbanes-Oxley. Each system has advantages and disadvantages and different business may use different systems (Fensterstock, 2005).

The neural network-based systems, is able to decide which characteristics are the most necessary to include in the prediction of credit risk. The basis of this model is to link how the brain, using a network of neurons, would process information. One of the main advantages of this model is that it can map out nonlinear relationships between the independent variables and the predictive variable. The most common model of neural network displays a multilayer perceptron. We see the bottom layer holds the input layers, such as the applicant's attributes, however these neurons does not automatically go to the output layer, credit score, but goes through a hidden layer. Thus the hidden layer receives information inputs from the previous layers. This hidden layer can also be thought of as the training phase in which information is provided and the weights are adjusted in order to produce a better output.

There is contradictory evidence of the effectiveness of neural networking for credit scoring. While Tam and Kian (1992) and Desay, Crook & Overstreet (1996) state that neural network is a better method, Altman et al. (1994) and West (2000) indicate that linear discriminate analysis and logistic regression, respectively, perform better and gives more accurate results. However, one huge drawback of this method is that it lacks explanation of how and why prediction was achieved. Bank managers using this system will not know the weights or relationships within the system. This makes it difficult to assess the model's decisions properly and accurately.

Statistical-based systems use multivariate regression models to approximate the probability that a customer will default on their loan. Unlike the judgmental system, weights assigned to factors are based on statistical assessment rather than human arbitration, thus this allows a statistical analyst to choose variables to check if a relationship makes financial sense. This also allows one to correct the accuracy of the model by finding sources of error and correcting them. However, this system requires someone with a background in statistics and can be hard to apply in some instances.

The last kind of system is the genetic algorithm-based system based on Charles Darwin's theory of "survival of the fittest" and may be the best model. Genetic algorithms (GA) create a random initial generation of models where the next generation is made up of the best of the first generation that have been tested for fitness against specific standards, which means an evolution of more advanced and accurate models. Unlike the other models, GAs can use 100% of the available data instead of just selected pieces. On the other hand, this model has been used very little, especially outside of universities and thus implementing a system that is not fully understood can be risky.

While there are many ways to estimate credit scores, credit agency does not divulge estimation of the credit score and thus considered a “black box.” Citron & Pasquale thus names three problems with the credit score system: 1) their opacity, 2) their arbitrary results, and 3) the disparate impact. Credit bureaus lack of transparency on their scoring methodology leaves individuals powerless to understand or challenge their score. Due to this opacity there exist arbitrary results. Different credit bureaus may present totally differing scores for the same individual. The secret behind the black box does not assure us of equal opportunity scoring. In fact, the scoring results show there is a disparate impact where women and minorities are concerned. Since credit score estimation is based on credit history alone, they fail to classify an entire group who may not have any credit history due to recent entrance into market, lack of large consumption in need of credit, or the fact that they rent instead of own assets. While credit score is a step in the right direction, the past financial crisis has shown that credit can still be given out incorrectly.

3. Methodology

Two common used methodology when working with probability of default is discriminant analysis and logistic regression. While both methods are used to analysis data with categorical outcomes, they do have different underlying assumptions. Linear discriminant analysis assumes normal distribution in the explanatory variables. Logistic regression, however, does distribution assumptions of the independent variables, thus it is more general.

Discriminant Analysis

The main objective of discriminant analysis is "to classify objects into one of two or more groups based on a set of features that describe the objects" (Gurny & Gurny 2013). In this case, we seek to classify good borrowers and bad borrowers based on different variables that describe that person. Thus the basic idea is to determine whether these different groups vary in means and if they can be used to predict default. Its primary use it to classify and make predictions where the dependent variable is in qualitative form and then find a linear combination which "best discriminates between the groups" (Altman, 1968). A disadvantage using discriminant analysis is their list of assumptions. Data is assumed to be normally distributed, variance and covariance are homogeneous, there is no correlation between means and variances, multicollinearity, and random sample. One advantage of using this method is that it reduces the space dimensionality by G-1, where G is the number of groups. In our paper, we only have two groups (Good or Bad) and thus we have one dimension.

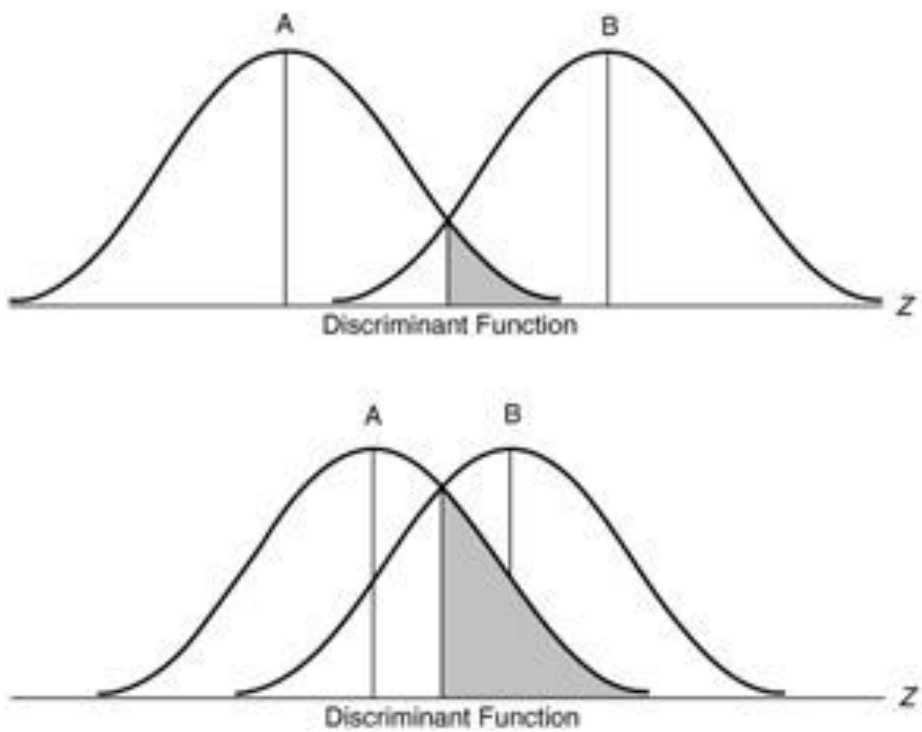
Discriminant analysis follows two basic steps. The first step is to estimate the coefficient of the independent variables, the borrower's characteristics. The coefficients serve as weights that measures which variables are good predictors for default. The second step is to apply a discriminant function to establish a cut-off value. The discriminant function is derived using the following equation:

$$Z = v_1 x_1 + v_2 x_2 + \cdots + v_n x_n \quad (1)$$

Where *v* is the discriminant coefficients and *x* is the independent variables. The discriminant function is treated as a standardized variable, so it has a mean of zero and a standard deviation of one. The discriminant coefficient maximizes the distances between the means of the dependent variables, where good predictor variables have the larger coefficient. Thus the discriminant function coefficient range between values of -1.0 and 1.0 and treated as a standardized variable. Thus the magnitude of the coefficients indicates the contribution of the independent variable.

<u>Figure 2: Discriminant Function distribution</u>

This is a hypothetical example of two groups (A, B). The discriminant function distribution diagram measures how well they are able to classify objects. If the overlap of the distribution function is small, than the function does a good job in classifying objects, if the overlap is large however there is a large probability of misclassification and thus the function is poor.



An individual's z-score can be found by simply summing the product of the coefficient with the independent variable. The group mean is the average of all the individual's score, also referred to as the centroid (Stamatis, 2003). The success of the function can be determined by measuring the group centroid distance from one another. Figure 2 illustrates an example of the distribution of scores of two group functions. The key to evaluating the function is by measuring the overlap of the distribution. Thus the top diagram depicts a statistically significant function and does well in distinguishing between group A and B. The bottom diagram, however, show that group have a large overlap, and thus the function has a high probability of misclassifying borrowers.

Logistic Regression

Since our objective is to find whether loan default will occur or not, than the appropriate methodology to apply would be logistic regression. Thus logistic regression takes a binary variable which only takes two values, zero or one. The main objective of a logistic regression is to find the best fitting model to describe the relationship between the dichotomous characteristic of interest (the dependent variable) and a set of independent variables. Logistic regression generates the coefficients of a formula to predict a logit transformation of the probability of the presence of the loan characteristics. A logit function thus stipulates the probability that default will occur and one minus this function specifies that default will not occur.

$$Score_i = \alpha + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_K x_{iK} \qquad (2)$$

$$z = \alpha + \sum \beta_n X_n \qquad (3)$$

This is a standard scoring model in which α is a constant and Xs are independent variables such as credit score, age, and other loan characteristics. In this paper since we seek to determine how credit score can find the probability of default of bank loans so our first equation will be to test this theory. Credit score ratings can be ranked in different groups taking into account 3734the approved loan amount.

$$f(z) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-(\alpha + \sum \beta_n X_n)}} \qquad (4)$$

The output of this equation, which should be between 1 and 0 reveals the riskiness of the bank. An output of 0 or close to zero means the bank has low risk while an output of 1 or close to one means the bank has high risk. This logistic function can be rewritten as a logistic model by using the expression of the probability of X. Logistic regression models the probability associated with each level of the response variable by finding a linear relationship between predictor variables and a link function of these probabilities. First we need to link it with our scores variables in which:

$$Prob(Default_i) = F(Score_i) \tag{5}$$

The logistic distribution function can then be written as

$$Prob(Default_i) = \frac{\exp(b'x_i)}{1 + exp(b'x_i)} = \frac{1}{1 + exp(-b'x_i)} \tag{6}$$

A very common way to estimate the weights of the coefficients is to use the maximum likelihood method. Maximum likelihood estimation is used and is the product of the sum of the logit function when default occurs multiplied by the product of the sum of one minus the logit function when the default does not happen. Then maximize the log of the likelihood function in order to find the weights:

$$(Y_i = 1) \rightarrow Prob\ (Default_i) = \Lambda(b'x_i)$$

$$(Y_i = 0) \rightarrow Prob\ (No\ Default_i) = 1 - \Lambda(b'x_i) \tag{7}$$

$$L_i = (\Lambda(b'x_i))^{y_i}(1 - \Lambda(b'x_i))^{1-y_i}$$

$$\ln L_i = \sum_{i=1}^{N} y_i\, ln(\Lambda(b'x_i)) + (1 - y_i) ln(1 - \Lambda(b'x_i)) \tag{8}$$

The logit model uses the logistic distribution function to link the variables. Two steps are required in order to find the coefficients: 1. Set first derivative to 0 and 2. Use the Newton's method.

$$\ln L_i = \sum_{i=1}^{N} y_i\, ln(\Lambda(b'x_i)) + (1 - y_i) ln(1 - \Lambda(b'x_i)) \tag{9}$$

(10)

(11)

1. $\frac{\partial lnL}{\partial b} = \sum_{i=1}^{N}(y_i - \Lambda(b'x_i))x_i$

2. $\frac{\partial^2 lnL}{\partial b \partial b'} = -\sum_{i=1}^{N} \Lambda(b'x_i)(1 - \Lambda(b'x_i))x_i x_i$'

$$b_1 = b_0 - \left[\frac{\partial^2 lnL}{\partial b_0 b \partial b'_0}\right]^{-1} \frac{\partial lnL}{\partial b_0}$$

Lawrence & Arshadi (1995), Campbell & Dietrich (1983), Gardner & Mills (1989) all use logit models to analyze loans, in fact, Charitou, Neophytou and Charalambous (2004) states that the logit method is superior when predicting defaults.

4. Data

The data that I will be using is from a local credit union from 2006 to December 2014. I will be looking at two different datasets: 1) current loan portfolio and 2) charge-off loans. Figure 3 depicts a comparison of the loan portfolio and the charge-off loans. As of December 2014, the loan portfolio was valued $297,466,374 while the charge off loans were at $147,850.07. So roughly .05% of their portfolio loans were charged off.

Table 1 shows the descriptive statistic of the dataset used. More information has been collected on the active loans compared to charge-off loans, we received 42,650 active loans. After cleaning up the data, there are 22,446 active loans. Information about the interest rate, original balance, current balance, loan maturity, the borrower's credit score, available credit, and loan description is given in this data set.

Figure 3: Charge-offs/Portfolio

The data used is from a local credit union. This graph shows the amount of the active loan portfolio balance and the charge-off during 2006 and end of 2014.

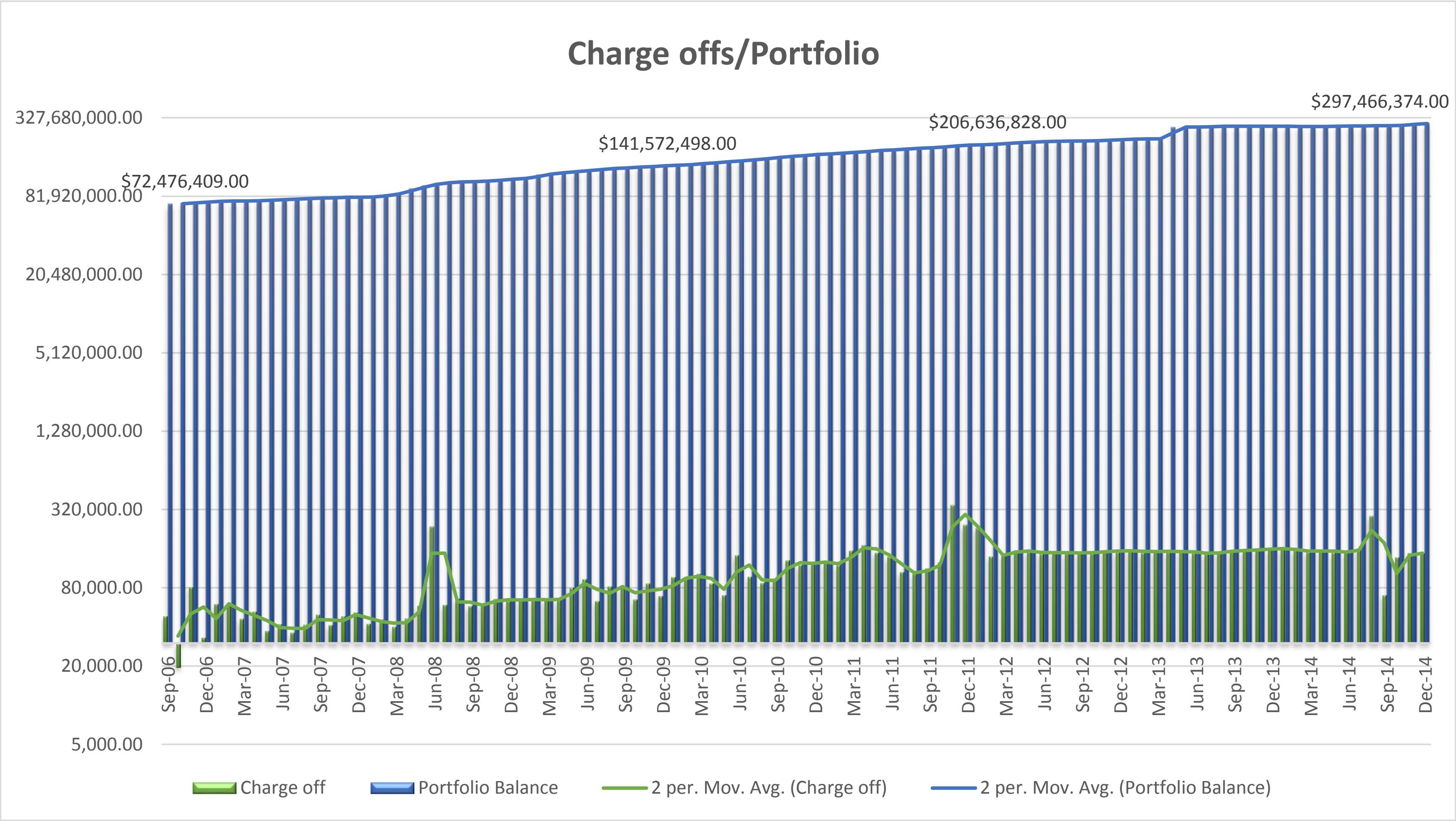


Table 1: Descriptive Statistics

Two different types of dataset was collected from local credit union from 2006 to 2014. Panel A is the summary statistics of the active loan portfolio. Current and Available balances were updated December 2014. Panel B is the summary of the charge off loans.

* Available credit, original balance and current balance is in the millions

Panel A: Active Loan Portfolio

Variable	Obs	Mean	Std. Dev	Min	Max
Interest Rate	22446	8.42	5.22	0	24
Credit Score	22446	609.18	241.22	0	964
Available Credit	22446	87136.25	364944.2	0	163.36*
Original Balance	22446	17558.55	36581.64	0	1.835*
Current Balance	22446	12873.98	46161.29	0	4.76*
Maturity	22446	4.38	6.21	0	36

Panel B: Charged-off Loans

Variable	Obs	Mean	Std. Dev	Min	Max
Credit Score	578	601.2	63.39	439	839
Age	577	41.39	12.68	21	83
Amount	578	3536.82	4065.38	3.2	29443.04
Debt Ratio	578	27.38	12.86	1.37	100
Delinquency	573	463.54	247.65	2	1002
Duration	571	2.08	1.68	0.01	13.67

The charge-off loan database is a much smaller dataset with 3,371 observations. After cleaning up the charge-off data, we were left with 578 observations which can also be seen in Table 1. The dataset also provides information about the borrower's age and credit score, and the loan amount, duration, and loan type. We deleted observations that had missing data for credit scores and debt ratio. Also some observations seemed to be mistyped (for example an individual had a 2006 credit score value) and those were also deleted. While having more variables would be optimal, this is a good starting point and since very few researchers have the availability of bank data this will provide great insight. Table 2 provides definition of the variables that will be used throughout this paper. As can clearly be seen both dataset provides different variables with the exception of credit score which in included in both.

Table 2: Variables Description

Active Loans	
Credit Score:	The borrower's credit score calculated by FICO indicates there risk based on credit history
Available Credit:	The remaining amount of an open line of credit or revolving loan (Credit line limit minus borrowed/used amount; since December 2014)
Original Balance:	The loan amount taken out by borrower at issuance date
Current Balance:	The amount of the loan the borrower still owes (original balance minus payments made; since December 2014)
Interest Rate:	The cost of borrowing
Maturity:	The life of the loan
Charge-Off	
Credit Score:	The borrower's credit score calculated by FICO indicates there risk based on credit history
Age:	Age of the principal borrower
Amount:	The amount of the loan charged-off
Debt ratio:	Total Debt/Total Assets
Duration:	The time that borrowers actually paid back the loan (date last paid minus issuance date)
Delinquency:	The time (in days) the borrower has not paid back their loans (since December 2014)

We then separated the default dataset by loan type and calculated the mean, standard deviation, and sample size for each individual loan type which is summarized in table 3. Personal loan, used auto loans, indirect loans and visa loans have the highest number of observation. For the rest of this paper we will focus on personal, visa and used, indirect, and new auto loans. Reasonably, auto loans have a larger loan amount then the entire data set (as well as credit score), while visa and personal loans have a lower loan amount. The data set as a whole is extremely volatile and randomly distributed. However, it is apparent that loan types with a larger sample size have a higher standard deviation in relation to the mean, while loans with a smaller sample size have smaller standard deviations in relation to the sample mean. Thus for the rest of this analysis we will be using the logarithmic of the variable to normalize the effects.

<u>Table 3: Charge-Off Descriptive Statistics by loan type</u>

	Obs	Amount Mean	Amount Std.Dev	Credit Score Mean	Credit Score Std.Dev	Debt Ratio	Debt Ratio Std.Dev
Personal	221	1630.66	1782.77	589.3	61.17	28.67	14.23
Used Auto	175	5207.39	4981.51	603.62	65.06	25.99	10.22
Indirect Auto	87	5135.66	3854.69	617.60	57.12	25.13	10.05
VISA	51	2443.78	2504.15	616.90	60.14	28.42	11.82
New Autos	21	7260.43	6583.52	606.38	50.69	25.98	13.72
Achiever	7	30.45	16.92	503.86	45.43	46.63	33.42
Line of Credit	7	3681.03	2838.68	645.71	75.04	30.90	16.50
Recreational Vehicle	5	3625.10	2722.53	670.20	64.58	23.57	13.02
Helping Hand	3	335.43	159.61	549	0	11.90	7.69
2nd Mortgage	1	3594.7	-	650	-	32.02	-

Table 4 indicate different characteristic of good borrowers and their loans. Good borrowers are defined as borrowers who have a ranking of at least a B. This ranking defines the original risk of the loan and is assigned by the bank manager. They are assigned a binary variable of 1, and thus all who have a binary of 0 are classified as bad borrowers. Not surprisingly, the mean of a good borrower has a credit score of 647 and that of a bad borrower is 498. Thus good borrowers have significantly higher credit score. Credit score is used to price interest rate for the borrowers. Thus it makes sense that bad borrowers, who tend to have lower credit scores, also has higher interest rates. Looking at original balances granted to borrowers, the table indicate that good borrowers have a significant higher loan amount, almost a $10,000 difference, then bad borrower. This is a reasonable deduction since banks doubt bad borrower capability to pay off a big loan while trusting good borrowers' ability to take on a bigger loan and pay it off. Available credit is the difference between the credit limit of a credit card and the amount already used. Thus available credit is the share of the line of credit that has not been spent. As for the case of original balance, the available credit is statistically significantly larger for good borrowers than bad borrowers. Good borrowers thus are granted a higher limit than bad borrowers. Maturity of the loan in this study can be defined as the length of the life of the loan. This was found by looking at the original date and the due date of the loan. Thus this figure shows that good borrower tend to have loans with longer maturity than bad loans. In order to reduce the risk of bad borrower loans they will give them a loan with shorter maturity.

This table also presents the different types of loans that this particular credit union gives to consumers. A binary variable was created to indicate what type of loans borrowers took out. The loans available are: auto, personal, share, credit card, end line credit, home equity line, trailer, second mortgaged, first mortgage, land, or business loan. Another possible explanation of

Table 4: Borrower and Loan Characteristics

Means of the borrower's loan characteristics and their significance. Using dummy variables to account the different type of loans given to borrowers. Where "Good Borrowers" were those borrowers classified as B risk or above. The rest were classified as "Bad Borrowers". Their means and differences are recorded in this table.

**** 1% Statistically Significant*
*** 5% Statistically Significant*
** 10% Statistically Significant*

Variable	Bad Borrower [0]	Good Borrower [1]	Difference[0]-[1]	T-statistics
Credit Score	498.33	647.66	-149.33	[-43.28]***
Interest Rate	14.22	6.40	7.82	[103.99]***
Original Balance	10326.65	20069.59	-9742.94	[-22.22]***
Current Balance	7829.79	14625.41	-6795.63	[-7.44]***
Available Credit	21604.09	109890.20	-88286.11	[-23.39]***
Maturity	3.79	4.59	-0.80	[-9.43]***

Variable	Bad Borrower [0]	Good Borrower [1]	Difference[0]-[1]	T-statistics
Auto Loan	0.416	0.451	-0.035	[-4.58]***
Personal Loan	0.249	0.122	0.127	[20.39]***
Share Secured Loan	0.024	0.012	0.012	[5.32]***
Credit Card	0.204	0.280	-0.076	[-11.99]***
End line of Credit	0.049	0.040	0.009	[2.85]***
Home Equity Line	0.007	0.009	-0.002	[-1.46]*
Trailer Loan	0.001	0.009	-0.008	[-9.66]***
Second Mortgage Loan	0.007	0.013	-0.006	[-3.94]***
First Mortgage Loan	0.009	0.041	-0.031	[-15.75]***
Business Loan	0.009	0.000	0.009	[7.17]***
Land Loan	0.003	0.004	-0.001	[-1.31]*

the larger loan amount for good borrowers is the fact that good borrowers tend to invest in more expensive tangible items such as autos, homes, and land. Table 4 shows that good borrower take out more auto loans, home equity lines, trailers, first and second mortgages, and more land loans. In addition to these investments, good borrowers have more credit cards. Bad borrowers, on the other hand, tend to take out more personal loans, share secured loan, open end line of credit and business loans compared to good borrowers.

Table 5 compares the means of the variables for the charge-off dataset categorized into loan type. For the age characteristic, it is not significant for any loan type except for achiever loans. The mean age for most charge-off borrowers are in the lower 40s. Therefore, the idea that younger borrowers are more likely to default on their payment is not substantiated by looking at just the means. The achiever loan is the only loan that is statistically significant and that its age is lower than 40s. Achiever loan borrowers are seeking to build credit and one type of borrower who lacks most in credit history are young adults who have not had the opportunity to build a history. The charge-off amount is the statistically significant for all the loan types. The highest charge-off amounts are from auto loans (new, used and indirect loans). Since borrowers take out larger loans to afford an auto, it stand to reason that they will have higher charge-off amounts. The achiever loan is the lowest charge-off amount. Other than the fact that there is only few observations, achiever loans by definition is a small loan with the sole purpose to build up credit. The credit score variable is significant for all loan types except for used and new auto loans. Indirect auto and credit card have the highest credit score. Indirect auto loans usually are originated in the car dealership and then transferred to the bank. Thus the bank does not have direct contact with the buyer. Credit card application, while many times dealt through the bank, also has a third party involved, the credit card company (in this case Visa). Thus these third party

Table 5: Charge-off Loan Characteristics

Dummy variables are assigned to different loan types in the charge-off dataset. The means of variable are than compared by specific loan versus the rest of the loans.

**** 1% Statistically Significant ** 5% Statistically Significant * 10% Statistically Significant*

Variable	Other Loans [0]	Personal Credits [1]	Difference[0]-[1]	T-statistics
Age	40.74	42.34	-1.60	[-1.461]
Amount	4716.82	1630.66	3086.16	[11.369]***
Credit Score	608.56	589.30	19.26	[3.621]***
Debt Ratio	26.58	28.67	-2.09	[-1.827]*
Delinquency	409.25	551.95	-142.70	[-6.886]***
Duration	2.65	1.14	1.51	[13.648]***
Variable	**Other Loans [0]**	**Credit Cards [1]**	**Difference[0]-[1]**	**T-statistics**
Age	41.24	42.51	-1.27	[-0.688]
Amount	3642.60	2443.78	1198.82	[3.04]***
Credit Score	599.68	616.90	-17.22	[-1.943]**
Debt Ratio	27.28	28.42	-1.14	[-0.651]
Delinquency	452.33	583.41	-131.08	[-3.796]***
Duration	1.96	3.36	-1.40	[-2.93]***
Variable	**Other Loans [0]**	**Indirect Auto [1]**	**Difference[0]-[1]**	**T-statistics**
Age	41.59	40.03	1.56	[0.989]
Amount	3253.52	5135.66	-1882.14	[-4.167]***
Credit Score	598.29	617.60	-19.31	[-2.851]***
Debt Ratio	27.78	25.13	2.65	[2.148]**
Delinquency	489.06	321.00	168.06	[6.923]***
Duration	2.04	2.30	-0.26	[-1.665]*
Variable	**Other Loans [0]**	**Used Auto [1]**	**Difference[0]-[1]**	**T-statistics**
Age	41.86	40.19	1.67	[1.491]
Amount	2811.39	5207.39	-2396.00	[-5.817]***
Credit Score	600.14	603.62	-3.48	[-0.598]
Debt Ratio	27.98	25.99	1.99	[1.920]**
Delinquency	487.95	408.02	79.93	[3.749]***
Duration	1.87	2.57	-0.70	[-5.188]***
Variable	**Other Loans [0]**	**New Auto [1]**	**Difference[0]-[1]**	**T-statistics**
Age	41.28	43.24	-1.96	[-0.854]
Amount	3396.43	7260.43	-3864.00	[-2.672]***
Credit Score	601.00	606.38	-5.38	[-0.473]
Debt Ratio	27.43	25.98	1.45	[0.476]
Delinquency	468.66	328.90	139.76	[3.054]***
Duration	2.03	3.41	-1.38	[-3.932]***
Variable	**Other Loans [0]**	**Achiever [1]**	**Difference[0]-[1]**	**T-statistics**
Age	41.47	31.71	9.76	[2.842]**
Amount	3579.81	30.45	3549.36	[20.816]***
Credit Score	602.39	503.86	98.53	[5.672]***
Debt Ratio	27.14	46.63	-19.49	[-1.542]
Delinquency	466.97	186.29	280.68	[5.443]***
Duration	2.10	0.35	1.75	[18.405]***

loan transactions may require a larger credit score cut-off before being accepted. Personal loans have a lower credit score than the other loans. Relationship banking may have influenced the acceptance of this loan application. An achiever loan has the lowest credit score. Credit cards and personal loans have higher payment delinquency than the other loan types. These loans, which also have lower charge-off payment (and loan amount), are given more time in delinquency until marked off the books. Duration is the amount of time that the borrowers paid off their loan before defaulting. The three auto loan and credit cards loans have higher duration mean the other loan type. Thus they had more loan payment periods than the other loans.

Using these two dataset, we will attempt to see how and if credit score is a good measure for default. While some charge-offs are due to the borrower's death or incarceration, most are due to bankruptcy, post repo, inability to pay, or just the refusal to pay. In order to run a test to see whether credit score can really predict the probability of default we need a database that has both default loans and paid-off loans. In our portfolio loan database we find which loans have paid of at least 99% of their loans back and we assign the binary variable of 0. We merge them with our defaulted loans, classified as 1, and delete any default that was due to death, prison, or any charge-off amount below $50 or have been delinquent in the last 30 days. Our new database has 1261 observations with 543 being default observations and 718 being paid-off loans. The credit score ranges from 437 to 850 with a mean of 677. The only shared variable is credit score and thus our regression will focus on the influence that credit score has as a predictor of probability of default.

5. Results

Multiple Linear Regression

The next step of this data analysis is to run a multiple regression analysis on the default dataset. The main idea of this paper is to find the probability of default. Even though we will not be able to ascertain this using multiple regression, what this regression analysis will allow us to do is see if the credit score is a legitimate predictor for charge-off amount. Table 6 shows the result of our multiple regressions. While we do not know the exact method that companies use to calculate credit score, at least two variables in our dataset seem to have a positive significance with credit score, debt ratio and age. We focus our regression on debt ratio and age since these are characteristics of the borrower opposed to characteristics of loans. Therefore, older borrowers with lower debt ratio has better credit score. This is logical since older borrowers have more credit history and have had time to build their credit. Having a high debt ratio is also synonymous to having high risk, thus it is puzzling why this value is not negative. However the r-square is very low.

The second regression output is where the dependent variable is the log of charge-off amount. Credit score is clearly positively significant with charge-off amount in all four equation. This seems to indicate that borrowers with high credit score has higher charge-off amounts. This maybe because borrowers with higher credit score are given higher loan amount, thus higher potential for larger charge-off amounts. However, these equations have a low r-square. The r-squared for regression model four is just 10%. This means the 10% of charge-off amount is represented by the four predictors that we used in this regression analysis. Ideally, you want to have an r-squared of at least 50% for a model to be considered legitimate. In conclusion, the r-squared of 10% for this model is extremely low and could not be used to accurately predict charge-off amount.

Table 6: Linear Regressions

Multiple regression on the default database. Panel A is a multiple linear regression where the dependent variable is log(credit score) and the independent variable is log (debt ratio) and log(age). The p-value is included to show significance. Panel B is a multiple regression where the log(charge-off amount) is the dependent variable. Different models are run with different variables, emphasis is on the significance of credit score. R-squared is also recorded.

**** 1% Statistically Significant ** 5% Statistically Significant * 10% Statistically Significant*

A. Dependent: Log Credit Score

	Coefficient	Std. Error	P-Value
Constant	2.686	0.025	0.001
logdebtratio	0.028	0.008	0.001
logage	0.032	0.014	0.023
N	578		
R^2	3.11%		

B. Dependent: Log Charge-off amount

	(1)	(2)	(3)	(4)
		-	-	-
Constant	-6.639***	6.550***	5.381***	5.254***
logcreditscore	3.559***	3.472***	3.025***	3.062***
logdebtratio		0.111	0.130	0.130
logduration			0.259***	0.262***
logage				-0.145
R^2	7.16%	7.35%	9.74%	9.83%

Discriminant Analysis Results

We run a discriminant analysis on the first dataset. Here are categorical group is good borrower which is determined by the original risk of the borrower. If they were assigned a B or better they are classified as good borrower (1), if they have a risk below this than they are classified as bad borrowers (0). Since there is only two different groups there will only be 1 function. First, we look at how well credit score can preditct good borrowers. Table 7 shows the outcome and that the discriminant function is significant. However, using just that one variables comes with misclassifications. Discriminant analysis uses the group means in order to discriminate betweens goups. As can be seen in the following table, the group means are relatively close to each other indicating an overlap in the distribution graph. With an overlap, this means that there is a higher tendency of classification errors. While this model was able to correctly classify about 89% of the good borrowers, it had more problems classifying bad borrowers. This model only classified 16.32% correctly, almost 84% were misclassified in group 1. Thus using this model, they would place a high porpotion of borrower in good borrower standing when they really belonged as bad borrower classification.

Table 7: Discriminant Analysis of Good Borrowers based on Credit Score

Discriminant analysis on the group good borrower is run with one predictor, credit score. The first section of is the canonical test which shows the number of functions and its significance. The second section looks at unstandardized and standardized canonical coefficient for the discriminant function and the canonical structure. The next section looks at the group means, the "centroids." And the last section looks at the classification. Whether the model is able to classify correctly the type of borrower.

Fcn	Canon. Corr.	Likelihood Rato	F	Prob>F
1	0.103	0.989	240.06	0.000
Ho:	this and smaller canon. Corr. Are zero			

	Unstandardized Canonical	Standardized Canonical	Canonical structure
logcreditscore	1.069	1	1
constant	-2.659		

Goodborrower	Group Means
0	-0.176
1	0.061

TRUE Goodborrower	Classified 0	Classified 1	Total
0	944 16.32%	4,841 83.68%	5,785 100%
1	1,857 11.15%	14,804 88.85%	16,661 100%

We decided to expand the model to include other variables given such as original balance, current balance, available credit, maturity, and interest rate. As can be seen in table 8, the function is significant. The unstandardized canonical coefficients are the parameters that are used to find the discriminant score. Thus the discriminant score for the average borrower is -0.71774, which is much better than the previous model (0.318). The standard canonical coefficient can be used to rank the importance of the variable. In this case, interest rate definitely is the most significant predictor.The correlation sturcture are latent discriminant loading variables. They represents the correlation between the predictor and the discriminant function. It can also be used to assess the importance of the variable. Usually a variable with a correlation of .3 or higher is desirable. Interest rate is the only variable that has a high correlation. The group means, or the centroid, of the group is the average of the sum of all the discriminant score of each individual of each group. The farther they are apart the better because it will be better able to descriminate per group. As we can see the group means are relatively far apart and this model does a much better job of classification. This model correctly classified good borrower by almost 91%. While the bad borrowers are still harder to classify (72%), it is much better than the previous model. However, 1,621 borrowers were still classified as good borrowers when they actually were bad borrowers. Therefore, in order to improve this model we need better variables. As could be seen in standardized canonical coefficient and the canonical structure variables were relatively low and therefore many of these variables are not the best predictors to use.

Table 8: Discriminant Analysis of Good Borrowers

Discriminant analysis on the group good borrower is run against several predictors: credit score, original balance, current balance, available credit, maturity, and interest rate. The first section of is the canonical test which shows the number of functions and its significance. The second section looks at unstandardized and standardized canonical coefficient for the discriminant function and the canonical structure. The next section looks at the group means, the "centroids." And the last section looks at the classification. Whether the model is able to classify correctly the type of borrower.

Fcn	Canon. Corr.	Likelihood Rato	F	Prob>F
1	0.713	0.492	359.2	0.000
Ho:	this and smaller canon. Corr. Are zero			

	Unstandardized Canonical	Standardized Canonical	Canonical structure
logcreditscore	-0.121	-0.113	-0.102
logoriginalbalance	0.021	0.015	-0.166
logcurrentbalance	-0.022	-0.025	-0.06
logavailablecredit	-0.210	-0.469	-0.129
logmaturity	0.463	0.172	-0.0073
intrate	0.276	1.090	0.854
constant	-1.963		

Goodborrower	Group Means
0	1.724
1	-0.599

TRUE Goodborrower	Classified 0	Classified 1	Total
0	4,164	1,621	5,785
	71.98%	28.02%	100%
1	1,549	15,112	16,661
	9.30%	90.70%	100%

The next test looks at our merged database. This will give us a clearer picture of whether credit score can predict the probability of default which is shown in Table 9. Since credit score is our only variable in this database we are constricted in using just this predictor. The function again is significant and because only one predictor is use the standardized canonical coefficient and structure is one. The unstandardized canonical coefficient is given in which we are able to find the discriminant score function: $D_i = -9.99 + 0.015 * X_i$ where i is each individual borrower and x is their credit score. The idea behind the discriminant score is to find the group means. The group means for paid-off loans is 0.835 and for default loans, -1.104. Credit score is able to correclty classify loans that will default (be paid-off) by 85% (81%). Thus this model is relatively good. A manager has a reasonable vindication to assign a cut-off score to avoid charge-offs. Thus a manager will be willing to accept a credit score which discriminant score is closer to group mean of 0.835. The group mean's credit score is approximately a credit score of 722, thus anything above that should clearly be accepted. The average credit score of this database is a 677, even though the descriminant score is below the group mean, the discriminant score of .165 is much closer to the group zero's mean than group one's mean. Thus a credit score of 677 should also be accepted.

Notice that there is a big difference in credit score capacity of predicting actual probability of default and good borrower classification that was given by the bank. While credit score was able to classify good borrowers relatively well, it was not able to classify bad borrowers that well. Credit score model tended to over-classify borrowers as good borrowers where the bank had determined them to be bad borrowers. Thus is can be assumed that managers uses other variables other than credit score to decide whether a borrower is good or bad. Though credit score does an admirable job at classifying loans probability of default, there is still room for improvement. It still has a 19% probability that it will classify a loan as a pay-off loan when it actually will default. Depending on the amount of charge-off, this can be a huge loss for any bank. In order to make this model more accurate, more variables will need to be added in order for the model to discriminate more efficiently.

<u>Table 9: Probability of Default using discriminant Analysis</u>

Discriminant analysis is run on a merged dataset of defaulted loans and paid-off loans. The defaulted loans were given the binary value of 0 and the paid off loans were given the binary number of 1. There is only one predictor value, credit score. The first section of is the canonical test which shows the number of functions and its significance. The second section looks at unstandardized and standardized canonical coefficient for the discriminant function and the canonical structure. The next section looks at the group means, the "centroids." And the last section looks at the classification. Whether the model is able to classify correctly the type of borrower.

Fcn	Canon. Corr.	Likelihood Rato	F	Prob>F
1	0.693	0.52	1162.8	0.000

Ho: This and smaller canon. Corr. Are zero

	Unstandardized Canonical	Standardized Canonical	Canonical structure
Credit score	0.015	1	1
constant	-9.99		

Default Loans	Group Means
0	0.835
1	-1.104

TRUE Default Loans	Classified 0	Classified 1	Total
0	580	138	718
	80.78%	19.22%	100%
1	79	464	543
	14.55%	85.45%	100%

Logistic regression

The credit union ranks their borrowers from an A+ to an E scale, A+ obviously being the safest borrower. While many banks classify good borrowers based on their credit score alone, using logistic regression, we can look at other characteristics that can help identify good borrowers as well. We run eight different regression with varying number of independent variables to see how that will change the degree of coefficients. The results can be seen in Table 10.

We are first interested in credit score. We want to know how well credit score predicts good borrowers. One unit increase in log(credit score) will produce an expected increase likelihood that the individual is a good borrower by .227 units. Because this is a logistic regression we need to transform coeffieicent to odds ratio which can simply be done by taking the coeeficient's exponential. Thus the odds ratio becomes 1.255. Thus each credit score unit increase is associated with a 25% odds of being a good borrower over a bad borrower. If we wanted to find the probability of the individual being a good borrower we can plug the coefficients in to the equation: $Y = \ln(\frac{p}{1-p})$= .503+.227*log(creditscore). If we use the mean as our x=609, than the total will be 1.135. Since we are looking for the probability we will have to transform that product as well: $p = (\frac{e^y}{e^y+1})$=75.68% that a borrower with a credit score of 609 is a good borrower. Thus credit score is a relatively good indicator of borrower status. By adding more variables however we can be more accurate. Using the coefficient output of equation 6 and the means of those variables we get a probability of 97% that the average borrower in this bank portfolio is a good borrower. However, the r-square is still problamatic. Using other variables, such as borrower's income and account may make the model fit better and thus increase in accuracy.

Table 10: Logistic Regression of Good Borrowers

Logistic regression on active loan dataset. Good borrowers' variable is a binary variable denoted as 1, and bad borrower is 0. We check to see how different loan variables influence good borrowers.

$$\text{prob}(goodborrower) = \alpha + \beta_1 \log(\text{credit score}) + \beta_2 \log(originalbalance) + \beta_3 \log(currentbalance) + \beta_4 \log(availalbecredit) + \beta_5 \log(maturity) + \ \beta_6 \text{interest rate} + +\beta_7 \text{indirect} + \beta_8 \text{coborrow} + \beta_9 multiple\ credit$$

**** 1% Statistically Significant ** 5% Statistically Significant * 10% Statistically Significant*

	1	2	3	4	5	6	7	8
logcreditscore	0.227***	0.118***	0.164***	0.391***	0.449***	0.508***	0.508***	0.506***
logoriginalbalance		0.443***	0.806***	1.024***	0.968***	-0.666***	-0.756***	-0.763***
logcurrentbalance			-0.338***	-0.066**	-0.036	0.100***	0.099***	0.876**
logavailablecredit				0.405***	0.390***	0.516***	0.527***	0.526***
logmaturity					0.339***	0.358***	0.351***	0.346***
intrate						-0.510***	-0.516***	-0.517
indirect							0.538***	0.652***
coborrow							0.233***	0.202***
multiplecredit								0.293***
constant	0.503***	-0.908***	-1.249***	-4.035***	-4.238***	5.993***	6.289***	6.257***
Pseudo R^2	0.87%	2.50%	3.38%	9.53%	9.64%	49.71%	49.94%	50.08%

Most variables had the expected signs. For example interest rate is negative and credit score is positive across the board. Good borrowers are characterized as having high credit score which is used to price loans. Thus, good borrowers will have lower interest rate (cost) than bad borrowers. Maturity and available credit is also positively significant. Good borrowers are given a higher available credit because they are trusted to keep up with the payments. Maturity is also longer. A possible reason for this is that many good borrowers may take out a big loan that requires a longer time to pay off. Other variables that came in positively significant are indirect variable, coborrow variable and multiplecredit variable.

An interesting result is that the original balance variable is positively significant with good borrower until interest rate is taken into account in which it becomes negatively significant. A possible explanation is that banks are normally more willing for good borrowers to have large loans. However, since large loans are seen as riskier they tend to have higher interest rates. A possible solution then is for good borrowers to take multiple smaller loans at lower interest rates. Current balance also had a change in signs. It started negatively significant but when the maturity variable was introducted, this switch the sign to positively significant. Since good borrowers tend to have longer loan maturity than bad borrowers, bad borrowers pay off their balances quicker.

The result of the logistic regression for the default dataset is seen in Table 11 and 12. The first table categorizes the default borrowers by their credit ranking and uses loan type as the main independent variables and loan characteristics are used as controle variables; also achievement loans are excluded to avoid a dummy trap. For borrowers who rank in the A tier, they were more likely to default on their credit card loans. There is no other statistically significant figure. Although A credit rating is seen as the best and E is the worse, all of these have failed to pay off their loans. Thus credit rating does not always signify that the borrower will not default.

Table 11: Credit Score Ratings and Loan types

Run logistic regression on charge-off dataset. The dependent variables are borrowers who categorized by credit score ranking where A have the highest credit scores and E has the lowest credit scores. The independent variables are the different loan types where the characteristic of loans (italicized) are used as control variables.

*[Indicate p-values] *** 1% Statistically Significant ** 5% Statistically Significant * 10% Statistically Significant*

Dependent Variable	A Tier	B Tier	C Tier	D Tier	E Tier
	(1)	(2)	(3)	(4)	(5)
Credit Card	2.059	0.894	-0.571	-0.622	-0.661
	[0.056]*	[0.301]	[0.386]	[0.315]	[0.241]
Personal Loan	1.173	0.542	-0.263	0.188	-0.614
	[0.268]	[0.490]	[0.631]	[0.709]	[0.196]
New Auto	0.380	1.375	-0.664	0.079	-0.644
	[.768]	[0.150]	[0.427]	[0.910]	[0.420]
Used Auto	0.480	0.392	-0.255	0.224	-0.175
	[0.654]	[0.633]	[0.652]	[0.669]	[0.726]
Indirect	1.339	0.529	-0.110	-0.200	-0.478
	[0.214]	[0.536]	[0.854]	[0.720]	[0.387]
Log(Age)$_t$	0.417	-1.861	-0.352	0.219	0.930
	[0.667]	[0.060]*	[0.677]	[0.760]	[0.227]
Log(Amount)$_t$	0.469	-0.091	0.083	0.173	-0.458
	[0.079]*	[0.691]	[0.686]	[0.323]	[0.011]**
Log(Duration)$_t$	0.260	0.181	0.241	0.022	-0.418
	[0.806]	[0.646]	[0.493]	[0.942]	[0.161]
Log(Delinquency)$_t$	-0.096	-0.700	-0.035	-0.369	1.429
	[0.806]	[0.035]**	[0.914]	[0.172]	[0.001]***
Constant	-4.942	2.575	-0.892	-0.914	-4.344
	[0.032]**	[0.189]	[0.605]	[0.535]	[0.009]***
N	571	571	571	571	571
Pseudo R^2	4.40%	2.91%	0.52%	1.43%	5.02%

Table 12 classifies the default borrowers by credit type. First we focus on credit score as the sole predictor. In this case, we are not looking so much at the probability of default since all of these borrowers have defaulted, but looking at what probability that they defaulted with a specific loan type. An average borrower with a credit score of 601 is more likely to default on a personal loan than any other loan (probability of 38%). Credit score seems to be a good predictor for personal loans with the expected sign. For indirect and credit card loans however there is a positive significance which seems to be contrary to common rationality. Credit score was also insignificant for used and new auto loans.

Table 12: Credit Score Ratings and Loan types

Logistic regression was run on the charge-off dataset from a local credit union. The dependent variable is the different types of loans and the loan characteristics are the independent variables. Panel A focuses on credit score, while panel B looks at the other variables as well.

$$\text{Credit Type} = \alpha + \beta_1 \log(\mathit{Credit\ Score}\ ) + \beta_2 \log(\mathit{age}) + \beta_3 \log(\mathit{amount}) + \beta_4(\mathit{Duration}) + \beta_4(\mathit{Delinquency}) + \beta_4(\mathit{Debt\ Ratio})$$

*[Indicate p-values] *** 1% Statistically Significant ** 5% Statistically Significant * 10% Statistically Significant*

Panel A:

Dependent Variable	Credit Card	Used Auto	New Auto	Indirect	Personal
	(1)	(2)	(3)	(4)	(5)
Log(Credit Score)	6.119	1.158	2.343	7.032	-6.97
	[0.057]*	[0.562]	[0.631]	[0.006]***	[0.001]***
Constant	-19.359	-4.049	-9.788	-21.292	18.862
	[0.031]**	[0.466]	0.471	[0.003]	[0.001]***
N	578	578	578	578	578
Pseudo R^2	1.05%	0.05%	0.13%	1.53%	1.69%

Panel B:

Dependent Variable	Credit Card	Used Auto	New Auto	Indirect	Personal
	(1)	(2)	(3)	(4)	(5)
Log(Credit Score)	6.689	-3.741	-2.899	3.882	-0.191
	[0.068]*	[0.107]	[0.633]	[0.172]	[0.938]
Log(Age)$_t$	0.194	-1.247	1.450	-1.454	2.053
	[0.894]	[0.109]	[0.439]	[0.137]	[0.013]**
Log(Amount)$_t$	-0.572	0.958	1.423	1.100	-1.116
	[0.034]**	[0.001]***	[0.010]***	[0.001]***	[0.001]***
Log(Duration)$_t$	1.378	1.866	2.683	0.565	-2.959
	[0.004]***	[0.001]***	[0.002]***	[0.152]	[0.001]***
Log(Delinquency)	2.841	-0.386	-0.929	-1.14	1.423
	[0.001]***	[0.158]	[0.055]*	[0.001]***	[0.001]***
Log(debt ratio)	0.264	-0.078	-0.444	-0.716	0.340
	[0.688]	[0.860]	[.716]	[0.205]	[0.448]
Constant	-27.676	9.036	-0.485	-10.213	-3.401
	[0.007]***	[0.154]	[0.975]	[0.192]	[0.613]
N	571	571	571	571	571
Pseudo R^2	8.99%	12.39%	16.32%	11.58%	27.26%

Next we look at all the predictors. Once we put all the other variables, credit score is only 10% statistically significant for credit card loans. Thus credit score is not a very good predictor for default. This demonstrates that while credit score can be a measure of a borrower's risk, other variables should be taken into account when measuring probability of default. Debt ratio is also insignificant for all five loan types. Even though it may seem logical to assume that a person that has a low debt ratio, and thus a lower porportion of debt to assets, would be able to pay loan and thus avoid default, this can not be used as a predictor.

Table 13: Logistic Regression of Probability of Default

Logistic regression on merge data set of default and paid-off loans. The default loans were given a binary value of 1, while the paid-off loans were given the value of 0. The probability of default is the dependent variable and credit score is the sole independent variable.

$$\text{prob}(Default) = \alpha + \beta_1 \text{Credit Score}$$

Prob	Coef	Std. Err.	P-Value
Credit Score	-0.024	0.001	0.001
Constant	15.923	0.852	0.001
N	1261		
Pseudo R^2	43.37%		

The last logistic regression run is on the merge data. The results can be found in table 13. Credit score variable is negatively significant with a coefficient of -0.024. Thus a one unit increase in credit score, will produce a .024 decrease likelihood that the loan will default. Looking at the odds ratio, a one unit increase in credit score is associated with a 2.4% odd of decrease probability of default. Thus the probability of default of the average credit score borrower in this dataset (creditscore=677) is 41.9%. Obviously, the higher the credit score the less probability of default. An 850 credit score borrower would have a 1.1% chance of default while a borrower with a 500 credit score would have a 98.1% probability of default. However, the Pseudo R^2 is only 43% so other variable should be taken into account.

Table 14: Result Summary for Probability of Default

Credit Score	Discriminant Score	Probability of Default.
500	-2.490	98.1%
592 *(DS mean)*	-1.110	84.7%
663	-0.045	50.3%
667	0.015	47.9%
722 *(DS mean)*	0.840	19.7%
735	1.035	15.2%
786	1.800	5.0%
850	2.760	1.1%

The discriminant score mean had a score of 722, which has a 19.7% probability of default. If a bank want to lower their probability of default they must choose a score that is close to discriminant score mean, since this function does a good job of classification. The greater the credit score the lower the probability of default as can be seen in Table 14. This function and probability of default model however can be improved by adding other variables.

6. Conclusion

Most of the literature related to probability of default has focused on the bond market. The consumer lending market has been gravely overlooked in research mainly due to data available. Consumer lending however is a huge market that should not be overlooked. Due to the recent financial crisis banks have suffered a massive loss on loans. This brings up an awareness of a need for an approach that banks can adopt to mitigate these losses and to measure credit risk more accurately. The objective of this paper is to find a simple model that small bankcs could use in order to find the probaiblity of default of their loans.

Credit score is a wonderful tool to measure riskiness of a borrower. Using discriminant analysis and logistic regression we found that credit score is a predictor of default. Using discriminant analysis a manager could assign a cut-off score that will reduce the likelihood of default. However, its opaqueness and the fact the borrowers with good score default shows that it should not be the sole predictor. While the credit score model had passable result in classifying loans, they still misclassified loans. This error could cause a bank to lose money. The logistic regression also had a low r-square which also causes question to the accuracy of the model. In order to have a better predicting model, more variable should be included.

Adding more variables to the model will help in the accuracy of the prediction. This gives practical use to bankers when making the decision of whether to accept a loan. While credit score does give good information, it should not be the sole factor when making this decision. This model can also be used for existing loan. If a bank knows the probability of default of their existing loan they will be forwarned therefore forearmed in case the worse were to occur. Having this information is essential for a bank to make the proper credit rationing and capital adequacy decisions.

7. References

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

Altman, E. I., & Suggitt, H. J. (2000). Default rates in the syndicated bank loan market: A mortality analysis. *Journal of banking & finance*, *24*(1), 229-253.

Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of banking & finance*, *18*(3), 505-529.

Altman, E. I., Gande, A., & Saunders, A. (2010). Bank debt versus bond debt: Evidence from secondary market prices. *Journal of Money, Credit and Banking*, *42*(4), 755-767.

Barry, P.J. "Modern Capital Management by Financial Institutions: Implication for Agricultural Lenders. *Agr.Rev.Financ.Rev.*61 (2001): 103-22

BIS Material, June 2006. "Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version." *Bank for International Settlement*.

BIS Material, June 2011. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems - Revised Version June 2011." *Bank for International Settlement*.

BIS Material, July 1988. "International Convergence of Capital Measurement and Capital Standards." *Bank for International Settlement*.

Black, F., & Cox, J. C. (1976). Valuing corporate securities: Some effects of bond indenture provisions. *The Journal of Finance*, *31*(2), 351-367.

Campbell, T. S., & Dietrich, J. K. (1983). The determinants of default on insured conventional residential mortgage loans. *The Journal of Finance*, *38*(5), 1569-1581.

Citron, D. K., & Pasquale, F. A. (2014). The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, *89*.

Charitou, A., Neophytou, E., & Charalambous, C. (2004). Predicting corporate failure: empirical evidence for the UK. *European Accounting Review*, *13*(3), 465-497.

Demyanyk, Y., & Hasan, I. (2010). Financial crises and bank failures: a review of prediction methods. *Omega*, *38*(5), 315-324.

Demyanyk, Y., & Van Hemert, O. (2011). Understanding the subprime mortgage crisis. *Review of Financial Studies*, *24*(6), 1848-1880.

Desai, V. S., Crook, J. N., & Overstreet, G. A. (1996). A comparison of neural networks and linear scoring models in the credit union environment. *European Journal of Operational Research*, *95*(1), 24-37.

Duffie D, Singleton K. 2003. *Credit Risk.* Princeton: Princeton Univ. Press

Featherstone, A. M., Roessler, L. M., & Barry, P. J. (2006). Determining the probability of default and risk-rating class for loans in the seventh farm credit district portfolio. *Applied Economic Perspectives and Policy*, *28*(1), 4-23.

Federal Reserve Economic Data-Fred (2015). "Net Charge-Offs On All Loans and Leases, All Commercial Banks." Federal Reserve Bank of St. Louis.

Fensterstock, Albert "Credit Scoring and the Next Step," Business Credit: Mar. 2005. p. 12.

Fensterstock, Albert "Credit Scoring Basics," Business Credit: Mar. 1 2003. p. 47-78.

Gardner, M. J., & Mills, D. L. (1989). Evaluating the likelihood of default on delinquent loans. *Financial Management*, 55-63.

Geske, R. (1977). The valuation of corporate liabilities as compound options. *Journal of Financial and Quantitative Analysis*, *12*(04), 541-552.

Gurný, P., & Gurný, M. (2013). Comparison of Credit Scoring Models on Probability of Default Estimation for Us Banks. *Prague Economic Papers*, *2013*(2), 163-181.

Ivashina, V., & Scharfstein, D. (2010). Bank lending during the financial crisis of 2008. *Journal of Financial economics*, *97*(3), 319-338.

Jacobson, T., & Roszbach, K. (2003). Bank lending policy, credit scoring and value-at-risk. *Journal of banking & finance*, *27*(4), 615-633.

Jarrow, R. A., & Turnbull, S. M. (1995). Pricing derivatives on financial securities subject to credit risk. *The journal of finance*, *50*(1), 53-85.

Jarrow, R. A., Lando, D., & Turnbull, S. M. (1997). A Markov model for the term structure of credit risk spreads. *Review of financial studies*, *10*(2), 481-523.

Jarrow, R. A. (2009). Credit risk models. *Annu. Rev. Financ. Econ.*, *1*(1), 37-68.

Jiménez, G., & Saurina, J. (2004). Collateral, type of lender and relationship banking as determinants of credit risk. *Journal of banking & Finance*, *28*(9), 2191-2212.

Keys, B., Mukherjee, T., Seru, A., & Vig, V. (2008). Securitization and screening: Evidence from subprime mortgage backed securities. *Quarterly Journal of Economics*, *125*(1).

Kwan, Simon. 2001. “Rising Junk Bond Yields: Liquidity or Credit Concerns?” *FRBSF Economic Letter* 2001-33 (November 16).

LaCour-Little, Michael, and Yanan Zhang. (2014). "Default Probability and Loss Given Default for Home Equity Loans." *Economics Working Paper. SpringerReference*. Office of the Comptroller of the Currency.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives research, 2*(2-3), 99-120.

Lawrence, E. C., & Arshadi, N. (1995). A multinomial logit analysis of problem loan resolution choices in banking. *Journal of Money, Credit and Banking*, 202-216.

Longstaff, F. A., & Schwartz, E. S. (1995). A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, *50*(3), 789-819.

Lopez, J.A. (2002). The Empirical Relationship between Average Asset Correlation, Firm Probability of Default and Asset Size. Working Paper, Federal Reserve Bank of San Francisco.

Mayer, C., Pence, K., & Sherlund, S. M. (2009). The rise in mortgage defaults. *The Journal of Economic Perspectives*, *23*(1), 27-50.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates*. *The Journal of Finance*, *29*(2), 449-470.

Mian, A., & Sufi, A. (2008). *The consequences of mortgage credit expansion: Evidence from the 2007 mortgage default crisis* (No. w13936). National Bureau of Economic Research.

Saunders, A., & Allen, L. (2010). *Credit risk management in and out of the financial crisis: New approaches to value at risk and other paradigms* (Vol. 528). John Wiley & Sons.

Smith, C. W., & Warner, J. B. (1979). On financial contracting: An analysis of bond covenants. *Journal of financial economics*, *7*(2), 117-161.

Stamatis, D. H. (2003). Six Sigma and Beyond Statistics and Probability, Volume III.

Tufano, P. (2009). Consumer finance. *Annu. Rev. Financ. Econ.*, *1*(1), 227-247.

West, D. (2000). Neural network credit scoring models. *Computers & Operations Research, 27*(11), 1131-1152.

Zhou, C. (2001). The term structure of credit spreads with jump risk. *Journal of Banking & Finance*, *25*(11), 2015-2040.

Chapter 2

Essay 2: Christian Mutual Fund Performance

1. Introduction

For centuries, religion has influenced the history of this world and help shape it into what it is now. Religion is a personal belief that adheres to a supreme authority in which faith is placed on. It is reasonable to believe that religion also influences a person's everyday decision. According to the Huffington post the top 3 beliefs are (1) Christianity with 2.3 billion, (2) Muslim with 1.5 billion, and (3) nonreligious or atheist with about 1 billion people. According to the Gallop poll in 2013, 39% of the United States population attend a religious service weekly, furthermore 56% of the population consider religion very important in their lives (Newport, 2013). This paper seeks to explore the intersection between religion and finance by focusing on the case of religiously affiliated mutual funds.

Since religion is an important factor when making decisions, there has been an increase in demand for religious investment. This has brought an increase in popularity for religious mutual funds. Conventionally, religious mutual funds have been relegated as just another subset of a group called Socially Responsible Mutual Funds. Socially responsible mutual funds seek to invest in only those firms that meet their specific criteria. Social Investment Forum defines "Sustainable and responsible investing (SRI) an investment discipline that considers environmental, social and corporate governance (ESG) criteria to generate long-term competitive financial returns and positive societal impact." SRI thus undergoes a rigorous screening process assuring them that they meet these criteria. In the same manner, religious mutual funds tend to shy away from investing into "sin stocks" such as alcohol, tobacco, and gambling. Moreover, Protestant funds tend to avoid investing in industries that promote pornography, abortion, or same-sex marriages (Peifer, 2011). Many Catholic funds also avoid industries that are environmental polluters, that has excessive executive packages, or those that have poor labor relations. The main question that this paper asks is whether religious mutual funds are different than SRI funds and how they perform compared to SRI funds and the market. There remains a doubt on how effective a religious mutual fund would be compared to a regular mutual fund or market index who is not constrained with eliminating undesirable investment.

Since religious people have an added constraint to their investment choices, mainly that investment cannot compromise their values, this may affect their diversification value. There has been an increasing amount of literature that looks at SRI mutual funds' performance versus a benchmark and there has even been some research looking into Islamic finance, this paper seeks to contribute to this area by focusing on Christian mutual funds. The different types of fund that this paper will look at are: Protestant Funds and Catholic Funds. This paper seeks to separate Christian funds with social responsible funds and analyze their performance. Christian mutual funds tend be smaller than SRI as well as have a propensity to screen against moral ambiguous stocks rather than social or environmental screening. Due to restrictive screening that restrains diversification benefits, Christian mutual funds performs worse than the market and SRI funds. They have similar pattern of conventional mutual funds to perform better (compared to the market) during recession however there performance is still lower than SRI funds. This paper also seeks to compare Protestant and Catholic funds. There are a larger number of Protestant funds which also lean towards a larger NAV. On the other hand, Catholic funds take more risk but have had a greater return historically than Protestant. Also their funds have a higher percentage of efficiency.

This paper differs from others that it looks at Christian funds. Past literature related to this topic has focused on either SRI funds or Islamic funds, none that I am aware of looks exclusively at Christian funds. Using conventional methods as well as the novel data envelopment analysis method, this paper will attempt to answer whether Christian mutual funds are different than SRI funds and if they suffer due to their higher moral standards.

2. Literature Review

July 1774, in Amsterdam an investment trust called *Eendragt Maak Magt* was created and established as the first mutual fund (Rouwenhorst, 2004). The first mutual fund invested primarily in bonds issued to banks, foreign governments and plantation loans to the West Indies. These Investment Trust were later introduced to the United States in the 1890s. Now there are more mutual funds in the United States than securities listed in the NYSE. The number of mutual funds owned by individuals has increased precipitously over the last few decades. Mutual funds have become a popular tool of investment since it allows the individual investor to pool their funds with others and thus have a diversified portfolio managed by a professional. According to the 2012 Census, at the time there were 7,581 mutual funds, in other words, 44 percent of the US household population owns mutual fund.

Mutual Fund Literature

Mutual funds are considered an attractive investment for the following reasons: customer service, low transaction cost, diversification and professional management (Gruber, 1996). With all these appealing qualities, their performance, in many researchers’ observation, is less than appealing. Malkiel (1995), Gruber (1996), Carhart (1997), and Fama and French (2010) have all found that mutual funds are not able to outperform passive benchmarks and in most cases underperform passive indices. Gruber (1996) found that while some managers may augment value to funds, the investment cost (transaction cost, fund expenses, and loading fees) charged to investors eradicates value. Carhart (1997) surmises that while top decile mutual funds perform well, on average, most funds underperform due to investment expenses. He also disinterred evidence of a one-year momentum effect. Last year's winners tend to have higher than average return the next year, yet this effect discontinues the years after. He surmises that investment costs have a direct negative effect on funds' performance. Mutual funds returns are more enticing when they are reported in gross terms, before incurring transaction cost and expense ratios. When measuring in net returns, few are able to produce benchmark returns while covering costs. Malkiel (1995) however found that mutual fund underperform even before costs are deducted. Fama and French (2010) stipulates that those few funds that does outperform, only some are due to managers' skill oppose to luck.

However, Grinblatt and Titman (1989, 1993) find that superior performance can occur due to managers' skill in choosing stocks. They observe over performance specifically among aggressive-growth and growth type funds as well as with funds with the smallest NAV. However, their evidence is based on gross returns. Wermers (2000) tries to resolve both sides of the debate by studying gross equity holdings and net return of mutual funds. He finds that there is an annual 2.3% difference between net and gross return. The gross returns outperform the market index by 1.3% a year, and when costs are included, the net fund returns underperform by 1% per year. Of the 2.3% difference, 1.6% can be explained by fund expenses and transaction costs, while the remaining is due to unproductive fund holding such as bonds and cash. Thus, while managers can add value with their stock-picking talent, this is neutralized by the cost and expenses that investors incur.

On average, most researchers acknowledge that the net return of mutual fund is negative. However, Glode (2008) and Kosowski (2006) identified precedence where mutual fund have exceptional performance. They found that while mutual fund does underperform during expansion periods, during a recession the risk-adjusted performance has been positive. In fact, fund managers are more active during bad states than good states. The difference between alphas in recession and expansion periods is about 3-5% per year. Thus, an acceptable elucidation of investors' continuance interest in mutual funds is due to their role as insurance against economic downturns.

Social Responsible Investing

SRI funds comprise only those firms that practice social responsibility such as environmental policies or charitable donations. This requirement may be a hindrance to firms. Literature shows two main camps of thought regarding social responsibility. One side argues that engaging in social responsibility is costly and thus is an economic disadvantage for competitive markets (Friedman, 1970, McWilliams & Siegel, 1997). They claim that the money used in complying with social responsibility requisite means sacrificing potential profit making projects and as a consequence a drop in their potential net income. They strongly argue that the manager's primary responsibility is to the shareholders, being "socially responsible" is in fact acting irresponsibly. Yet it has been stipulated by a second group of researchers that engaging into social responsibility is actually a competitive advantage since it attracts investors, resources, quality employees, customers, and creates other unforeseen opportunities (Cochran & Wood, 1984; Waddock & Graves, 1997; Greening & Turban, 2000). Therefore, by engaging in social responsibility the firms are in essence receiving some good publicity that in the long run will be beneficial for them.

Although social responsible investing has grown and it keeps the investors feeling better about themselves and the work their money is doing; does it perform better than the conventional mutual funds? There has been a number of papers that look at this topic yet the result still seems unresolved. Guerard (1997), Diltz (1995), and Hamilton et al. (1993) found that SRI perform as well as benchmark portfolios. In other words, even though SRI firms had restricted investment, there is no real benefit to holding an SRI portfolio, but there is no harm either. Statman (2000) and Bartolomeo & Kurtz (1999), however, found that SRI perform better than normal portfolios. SRI funds benefit from their intense screening process which effectively eliminates the poor performing firm's thus resulting better performance than the benchmark. Yet another group of researchers found that SRI perform worse than the benchmark (Geczy, Stambough, and Levin 2003). Since SRI investment adds a constraint to the investment choices for their portfolio thus excluding not only certain stocks, but many times whole industry, this decreases the diversification benefits of the fund thus explaining the worse performance (Goldreyer and Diltz 1999). Barnett & Saloman (2006) tries to bring these differing views into accordance by explaining that when a fund undergoes intense screening then it will result in over-performance, since it eliminates poor firms, on the other hand, if it does very little screening, then it may still over-perform because their funds tend to exhibit more diversification. Yet those funds that are stuck in the middle with their screening process does not tend to do well. As can be seen through literature, these mixed reviews show that the performance of SRI is still unsettled.

Religious Funds

There is another subcategory of the SRI fund, which is the religious funds. Forte & Miglietta (2007) found that though religious fund are seen as a subcategory of SRI funds they should be seen as something different. Religious funds or "faith-based funds" show different characteristics from SRI funds that are easily distinguishable. Religious funds thus are distinguished by the values they adhere, there asset allocation, risk, and econometric profile. The investment strategy of a religious fund then is to not only stay away from “Sin stocks” but to invest in what their religion holds true. However, Miglietta’s paper uses Islamic funds to define religious funds and ignores Christian funds all-together.

However both Islamic and Christian funds tend to stay away from “sin” stocks. So how does neglecting sin stocks effect religious mutual funds? Sin stocks outperform Shariah-compliant stocks during both expansionary and contractionary economic periods (Liston & Soydemir, 2010). Hong & Kacperczyk (2009) looks at the effects of social norms on the price of sin stocks. Social norms indicate that normally individuals are against supporting companies that promotes human vices such as alcohol, tobacco and gaming. They also found evidence that “sin stock” are held less by institutional ownership and are also not covered much by analyst. Since they are neglected they tend to be cheaper than their counterpart and thus the market price is below the intrinsic price. Another reason why “sin stocks” are underpriced is due to litigation risk, these companies are perceived to be under constant regulatory scrutiny and thus their value is derived from very conservative accounting. Thus they found statistical significant evidence that sin stocks have higher expected returns both in the US and European markets.

Recently, many researchers have hone their focus towards Islamic financing and Islamic mutual funds. Islamic funds differ from conventional funds because they are Sharia-compliant and thus have limitation to their investment and cannot receive or pay interest (Kraeussl & Hayat, 2008). Thus they are not only prohibited in investing in companies that are against their values such as alcohol, gambling, pornography, entertainment and pork related industries but in addition to these industries they are also prohibited in investing in companies that deal with interest payment. Islamic funds tend to prefer small cap firms (Hoepner, Rammal & Rezec, 2009; and Girard & Hassan, 2008). This preference for small-cap is presumptively due to the reasoning that large cap companies may dabble in Sharia prohibited activities. Khatkhatay & Nisar (2008) suggests that funds are too liberal with their investment and needs to be more strict and restrictive of companies allowed to be part of their fund. Rubio, Hassan & Merdad (2012), Girard & Hassan (2006, 2008), Hussein (2007), and Kraussl & Hayat (2008) found that Islamic funds do not lose efficiency and are an effective investment opportunity. In other words, although Islamic fund must adhere to strict Islamic law, there is no difference in performance. Therefore, Islamic funds are a good alternative for investing while still upholding to Sharia law. However, Hoepner, Rammal & Rezec (2009) found that Islamic mutual funds trail behind the benchmark mainly due to its restrictions and prohibition of lucrative investment. Hussein (2007) finds that Shariah compliant indices underperform in a bear market yet that in a bull market they are a superior investment choice. However, Kraeussl & Hayat (2008) finds that equity funds tend to outperform the benchmark during a bear market. Hoepner, Rammal & Rezec (2001) also studied Islamic mutual funds in twenty different countries. Their finding are that they cannot determine whether Islamic fund generally under or over-perform, yet that national characteristics are important to explain fund performance. Interestingly they found that those funds located in countries where the largest Islamic financial centers were located tended to perform competitively and even outperform international equity market benchmark. On the other hand, those funds located in countries with less developed Islamic financial services or where the predominant religion was Christianity, tended to underperform their benchmark.

While religion in the financial, economic world has been neglected, it has not been all together ignored. Many researchers have come to the conclusion that religion is an important factor in an individual's decision process. There has been a common misconception that as science and academic learning advances in today's culture, then religion will cease to exist, yet Iannaccone (1998) observes that "the resurgence of evangelical Christians in USA, the rise of Islamic fundamentalism in the Middle East, and the explosive growth of Protestants in Latin America. In the US, there shows little or no decline of religion over time." Thus religion is an important factor in an individual's decision making and should not be ignored. In fact, Stulz (2003) uses religion as a prediction of culture. There have been many different philosophies on religion and economic performance. John Wesley, the founder of the Methodist church, was inclined to believe that Protestant ethics would bolster up economic development. Arruñada (2010) finds that "Protestant values shape a type of individual who exerts greater effort in mutual social control, supports institutions more and more critically is less bound to close circles of family and friends, and holds more homogenous values." It has been observed that Protestant countries tend to have more wealth and power than Catholic nations. One reason is that Protestant countries have better education (Becker & Woessmann, 2009). Max Weber (1904) attributes this to Protestant work ethics and their influence in developing capitalism. Protestants are taught that their hard work glorifies God. United States was founded with this Protestant work ethics as well. In early colonial times, John Smith admonished the residence of Jamestown for being idle and only letting the few work hard to maintain them, he thus quoted Paul saying "he that will not work, shall not eat," this ideology has continued and become a legacy of Protestant workers. However, this study does ignore that in many "Protestant" countries, religiosity is declining (for example European countries). Financial literature ascertains that an accession of religious participation has a statistically negative effect on economic growth and reduce individual's income (Barro & McCleary 2003;McCleary 2008; Lipford & Tollison 2003).

As individuals are more involved in religious activity they are predisposed to spend less time working and thus hampers economic development. However, while McCleary makes this case for several countries, United States is an exception and he also consented that Christian ethics is important to a child's upbringing.

Christian ethics are important and useful, McGuire, Omer & Sharp (2011) show that in those areas where religion is dominant, there are lower incidence of financial reporting irregularities. There is further indication that religion is apt to make individual and firm more risk averse, thus the portfolio structure may be more conservative in nature (Hilary & Hui, 2009). Adhikari & Agrawal (2014) also finds that banks headquartered in highly religious area takes less risk. Interestingly enough, of the two major Christian factors, Catholics exhibit less aversion to speculative risk than Protestants (Shu, Sulaeman, & Yeung, 2012). Kumar, Page & Spalt (2011) also observes this tendency that Catholics have higher pension for gambling and speculation than Protestants in that regions that have a higher Catholic ratio hold stocks with lottery-type features. Furthermore, they argue that religious beliefs do influence both individual's portfolio choice as well as corporate decisions. Thus it is important to quantify the importance that religion has on finance. This paper will try to fill in gaps that have not been looked at, mainly in the area of Christian mutual fund performance.

3. Data

The number of mutual funds owned by individuals has increased precipitously over the last few decades. Mutual funds have become a popular tool of investment since it allows the individual investor to pool their funds with others and thus have a diversified portfolio managed by a professional. According to the 2012 Census, at the time there were 7,581 mutual funds, in other words, 44% percent of US household population owns mutual fund. A proportion of this has been due to the growing number of interest to investing in line with their values. Using public data, we identify 111 Christian mutual funds and 153 SRI fund that span from January 2005 to January 2015.

Data was collected through publicly available websites. Social Responsible funds were collected from Social Funds website which is “the largest personal finance site devoted to socially responsible investing” and the USSIF, the Forum for Sustainable and Responsible Investment site. The Fama and French factors along with the market return and the risk-free rate was taken from French’s website. The descriptive statistics of Christian mutual funds and SRI fund returns are shown in Table 1.

By analyzing the descriptive statistics Social Responsible funds has a higher mean return than Christian mutual funds but also a larger standard deviation. The table depicts that SRI funds have a mean of .297% return with a considerably high standard deviation of 4.6%. The SRI fund fluctuated between a return of -38.6% and 31.3%. The worse return year was at the height of the financial crisis, 2008, but recovered substantially in 2009. On the other hand, Christian mutual funds has a statistically significantly lower mean of .15% return with a slightly smaller standard deviation of 4.3%. While the Christian mutual fund does have a wide range of fluctuation throughout the 10 years, the minimum return is lower than the Social responsible minimum return. This intimates that by having higher restrictions, Christian mutual funds are essentially “shooting themselves in the foot.”

Out of the two Christian sects, Catholic returns seemed to be slightly higher with a .28% mean return versus Protestant .13% mean return. However Protestant returns does have a slightly smaller standard deviation. In fact, Catholics have a higher return than the average of the whole CMF set, while Protestant funds has a lower mean. This seems to suggest that Catholics are

Table 1: Descriptive Statistics

	CMF Returns						SRI Returns			
	Mean	Std. Dev	Minimum	Maximum			Mean	Std. Dev	Minimum	Maximum
CMF	0.1517	4.3001	-31.0461	31.6076		**SRI**	0.297	4.6177	-38.5835	31.3131
2005	0.1656	2.701	-13.8322	6.5693		2005	0.2145	2.7161	-18.6373	26.4765
2006	0.3713	2.8991	-18.9593	12.8866		2006	0.6599	2.7964	-10.1739	14.3301
2007	-0.1672	3.5263	-19.5041	9.5904		2007	0.0997	3.3839	-22.427	14.8624
2008	-3.237	6.1919	-31.0461	12.4197		2008	-3.3376	6.9511	-38.5835	17.2227
2009	1.939	5.6791	-20.2299	31.6076		2009	2.2171	5.9419	-15.8898	31.3131
2010	1.0581	4.5562	-11.5691	13.1827		2010	0.9876	5.1835	-17.1769	19.7979
2011	-0.1282	4.4136	-26.0001	16.7283		2011	-0.415	4.8321	-27.4419	17.5376
2012	0.7181	3.0477	-11.6026	7.9826		2012	0.8817	3.29	-17.3285	16.94
2013	0.8501	3.1678	-25.665	8.4974		2013	1.324	3.1954	-14.1871	16.9872
2014	-0.1996	3.4106	-22.2322	8.5627		2014	0.0102	3.4595	-29.9556	18.2813
2015	-1.1782	2.1903	-7.7686	4.9727		2015	-1.365	2.1891	-6.016	7.7245

	Protestant Return						Catholic Returns			
	Mean	Std. Dev	Minimum	Maximum			Mean	Std. Dev	Minimum	Maximum
Prot	0.1292	4.2896	-31.0461	31.6076		**Cath**	0.284	4.3602	-22.6387	16.7617
2005	0.1512	2.6872	-13.8322	6.5693		2005	0.2725	2.8167	-6.015	6.4
2006	0.314	2.9131	-18.9593	8.3051		2006	0.7668	2.7836	-7.0752	12.8866
2007	-0.1741	3.5988	-19.5041	9.5904		2007	-0.1169	2.9569	-11.2092	4.7814
2008	-3.2989	6.1463	-31.0461	12.4197		2008	-2.7901	6.5211	-22.6387	6.1867
2009	1.8909	5.6444	-20.2299	31.6076		2009	2.2773	5.9308	-11.3452	16.7617
2010	1.007	4.519	-11.5691	12.8307		2010	1.3563	4.7716	-8.3976	13.1827
2011	-0.165	4.4027	-26	16.7283		2011	0.0676	4.4788	-11.2637	15.9517
2012	0.7388	2.9947	-11.6026	7.6179		2012	0.6173	3.2989	-11.2235	7.9826
2013	0.8402	3.1634	-25.665	8.4974		2013	0.8985	3.1961	-8.9633	7.1826
2014	-0.1519	3.2982	-17.7754	8.5627		2014	-0.4457	3.94	-22.2323	6.1611
2015	-0.8348	1.958	-4.5519	4.9727		2015	-2.9522	2.5133	-7.7686	0.5381

taking more risk in their investment but are having higher rewards compared to Protestants funds. Data of Thrivent mutual funds have also been collected. Thrivent financial is a Lutheran financial organization that has the largest Christian fund, however they are not included in the CMF dataset since they do not actively screen for or against anything. Compared to the other Protestant funds that does screen, they not only have a higher mean, but also a slightly lower standard deviation. Thus this seems to agree with the basic concept that more diversification reduces risk. Comparing the screen dataset versus the S&P 500 index, the returns of both the Christian mutual funds and Social Responsible funds is lower than the Index and the standard deviation is higher.

Christian Fund Data

When looking at Christian mutual funds, there are two different subsets: Protestants and Catholics. Though they have the same origins, the term catholic stems from the idea that there was one universal church yet that obviously changed with the Protestant movement. These two sects have a rich history of rivalry and although they have similar origins, there are some differing theological beliefs. However the Protestant sect also has different denominations that have different levels of conservatism, for example southern Baptist is known for their conservative behavior and strictly oppose alcohol, Lutherans, on the other hand, are more acceptable of the intake of alcohol. This paper will focus on these two sects of Christianity taking into account they contain different ethical values as well as risk tolerance. The specific funds that will be looked at are: from the Protestant Family: Praxis fund, New Covenant fund, Guidestone fund, Steward Fund and Timothy Plan fund. From the Catholic family: Ave Maria fund, Epiphany funds, and LKCM Aquinas fund.

Fund Background

Everence Financial service (Praxis) is a ministry of the Mennonite Church USA, started in 1945 which offers many different types of services for individuals, organization and congregations. Everence screening process adheres to their stewardship investing core values: respect, build, demonstrate, exhibit, support, and practice. Everence thus invests in companies that respect human right, ethnic and cultural diversity. They shun any company that promotes violence, such as weapons production and military contracting. They invest in countries that conduct in equal opportunity and fair compensation to their employees as well as companies that have sound corporate governance. They also positively screen those companies that support and develop their communities with their own resources. And lastly, they invest in companies that promote natural and environmental welfare.

New Covenant Funds are a part of the Presbyterian foundation group that makes their investment decisions which are consistent with views adopted by the General Assembly of the Presbyterian Church. The screening process avoids the gambling, alcohol and firearm industry. They also have positive screening for companies that hold fair treatment to employees and invests in their communities. However New Covenant does make a note that they may at times invest in companies that has been recognized as being in conflict with the principles held by the Presbyterian Church.

Guidestone, founded August 2001, belongs to the Southern Baptist denomination and is United States' largest screened Christian fund. This fund is much more socially conservative than the last few funds. Their goal is to provide high-quality, comprehensive mutual fund for individuals, foundations, retirement and other investment causes while adhering to their Christian values. They have an intense screening process against those companies that deals with liquor, tobacco, gambling, pornography, the abortion industries or any company who are irreconcilable to the Christian moral and ethics that Guidestone adopts.

Steward mutual fund is a non-denominational Christian fund that tries to remain consistent with their Christian faith. They only execute avoidance screening and thus do not have any positive screening, which promotes a typical Christian culture. The negative screening they apply is against abortion, alcohol, gambling, tobacco and pornography industries or those companies that derive significant income from these products.

The Timothy Fund was begun in 1992 by Arthur Ally. His purpose was to implement a fund that will more properly screen based on Christian value. Their goal is to be good stewards of what God has entrusted them. They mainly screen against 7 activities: abortion, pornography, entertainment that promotes violence and sexual immorality, alternative lifestyle, alcohol, tobacco and gambling.

The Ave Maria mutual fund is the most prominent catholic-oriented mutual fund. They hire Schwartz Investment Counsel, Inc. as adviser for their mutual fund. Their mission is to provide superior financial services while keeping with the Catholic teachings. Ave Maria holds on to a pro-life, pro-family philosophy. Their moral screening identifies companies and determines whether they are compliant towards the Catholic Church's values. These values mostly regard teachings on abortion, pornography, and policies that undermine the holy sacrament of marriage.

The Epiphany Fund is a Catholic fund which seeks to invest in securities that are consistent with Christian morals and ethical principles. They use the FFV scorecard (Faith and Family Values) to screen their investments. The objective of the FFV scorecard is to identify companies that are in keeping of their 4 pillars: 1) Life and Family; 2) Social Justice; 3) Environment; and

4) Corporate Governance. Their screening is consistent with the United States Conference of Catholic Bishops (USCCB) Socially Responsible Investment guidelines.

Luther King Capital Management (LKCM) Aquinas was founded in 1979 as an advisory firm committed to select equity based on Catholic values. There are committed in providing a solid financial performance while keeping in line with the Catholic values. They also follow the investment guidelines set by the USCCB investment guidelines. They screen against companies that engage in abortion, embryonic stem cell research and weapons of mass destruction. Along with this moral screening they also screen against companies that have poor environmental, human right records and employment records. This fund, however not only screens against companies, but also take a proactive stand discussing with companies about their practices that may come into conflict with their guidelines.

So while there is an obvious difference between Islamic funds and SRI funds do to not only restriction in “sin stocks, Islamic funds are very different because they restrict interest-bearing stocks. However these are questions on whether Christian funds are different than SRI funds. Both type of funds have similar screening quality which is why Christian funds have been categorized as a SRI funds. However, there Screening focus is different which impacts fund composition. As can be seen from table 2, CMF and SRI funds have different focus when it comes to screening. 93% of SRI funds have some sort of environmental screening. The second most popular screen is social screening. Social screening is a very broad subject that relates to anything that involves the improvement of society such as, but not limited to, community investment, human rights, and labor issues. On the other hand, they do not emphasize moral-type screening like abortion (only 1%) and pornography (6%). Christian mutual funds on the other hand accentuate the need to screen based on these issues that are very important to their faith, such as screening against pornography and abortion (96%). Another category that is also screening against by both CMF (83%) and SRI (69%) funds is the traditional exclusionary screens such as tobacco, alcohol, and gambling. Thus there is an obvious difference in screening emphasis on Christian mutual funds and SRI funds. An interesting matter to note is that 100% of the protestant funds execute traditional exclusionary screening process, while none of the Catholic funds do.

Table 2: Qualitative Screening

The data includes 108 Social Responsible Funds, and 106 Christian mutual funds. The information was collected from the fund family website or through the Social Investment fund forum. This table classifies what type of screening funds emphasizes on.

Screens	SRI	CMF
Environment	93%	25%
Governance	80	15
Social	84	49
Traditional Exclusionary	69	83
Weapons &Nuclear Power	69	25
other	39	-
Pornography	6	96
Abortion	1	96
Pharmaceuticals	5	17
Adult entertainment	-	34

Thus, at least qualitatively, we see a difference in screening undertaken by Christians compared to SRI funds through their screening process. Also according to the descriptive statistics Christian mutual funds has a lower standard deviation and thus their fund has lower risk. Thus values have an impact on screening which has an impact on fund composition and performance.

4. Methodology

Existing literature poses questions while analyzing Christian mutual funds. How does Christian mutual fund perform compared to the market? Is there a difference in performance during recessions as indicated by literature? Is Christian mutual fund performance different than SRI funds? How do Catholic and Protestant fund perform? We will be using traditional methods used in past literature to evaluate and compare performance.

Sharpe Ratio

When comparing performance of mutual funds, two popular ratios are the Sharpe ratio (1966) and Treynor ratio (1965). These two ratios are similar in theory and practice. The numerator is calculated by finding the excess return, the portfolio return subtracted by the risk free rate, divided by the standard deviation.

$$Sharpe\ Ratio = (R_p - R_f)/\sigma \tag{1}$$

where $R_p$the return of the portfolio is, R_f is the return of a risk-free asset, and σ is the standard deviation of the portfolio. Thus the Sharpe ratio calculates the excess return per unit of risk. The higher the Sharpe ratio the better since it indicates that the portfolio has performed well relative to the risk. However, if this ratio becomes negative it indicates that the investment in the portfolio is not worth the risk thus a risk-less alternative is preferable.

Sharpe (1994) details the importance of defining the differential return. $D_t = R_{Ft} - R_{Bt}$. He denotes that the differential return in period t is equal to the difference of the return on the fund in period t, R_{Ft}, and the return on the benchmark portfolio in period t, R_{Bt}, (in this analysis we will use both the risk free rate as well as the S&P 500 Index as the benchmark). The ex post

Sharpe Ratio than becomes $S_h = \frac{\bar{D}}{\sigma_D}$, where $\bar{D}$ is the average value of the differential return. The T-statistic can easily be found to measure the significance of the Sharpe Ratio: T-stat= $S_h * \sqrt{t}$.

Jensen's alpha and CAPM

Another customary measure used is the Jensen's alpha (1968).

$$\text{Jensen's alpha} = \propto = R_p - [R_f + \beta(R_M - R_f)] \quad (2)$$

where $R_p$the return of the portfolio is, R_f is the return of a risk-free asset, and R_M is the market return. This is another performance measure that presents the abnormal return of portfolio over the theoretical expected return given by the capital asset pricing model. Jensen's alpha is thus calculated by the capital asset pricing model (CAPM):

$$(R_p - R_f) = \propto \ +\beta(R_M - R_f) \quad (3)$$

The alpha in this model is the y-axis intercept of the excess return and thus signify a type of active return. If the value is positive, then it signifies that the portfolio is earning an excess return thus outperforming the benchmark. Thus a positive alpha is positive news since it indicates that the portfolio "beats the market."

Fama-French Factor model

While our focus remain on alpha as a measurement of performance, the Fama – French three factor model and the Carhart four-factor model are calculated in order to test for robustness. Therefore, we want to know to know the significance of alpha controlling for the additional risk factors and momentum.

The Fama and French (1993) three factor model is an extension of the CAPM model:

$$\left(R_p - R_f\right) = \propto\ +\beta_{i,1,t}\left(R_{M,t} - R_{f,t}\right) +\ \beta_{S,i,}SMB_t +\ \beta_{H,i}HML_t + \varepsilon_{i,t} \qquad (4)$$

where (R_p- R_f) is the excess return of the portfolio , $\left(R_{M,t} - R_{f,t}\right)$ is the excess return of the market which is measure using the value-weighted return of firms listed in CRSP as the return on market minus the one-month Treasury bill rate. The SMB_tvariable is Fama and French's "small minus big" factor which is a size loading factor and takes the three smallest portfolio minus the three biggest portfolio. HML is the "high minus low" factor which is a value loading factor that takes two value portfolio and subtract it by two growth portfolios.

We will also use Carhart four-factor model, which is an extended version of the Fama-French three factor model and includes a momentum factor.

$$\left(R_p - R_f\right) = \propto\ +\beta_{i,1,t}\left(R_{M,t} - R_{f,t}\right) +\ \beta_{S,i,}SMB_t +\ \beta_{H,i}HML_t + \beta_{M,i}MOM_t \qquad (5)$$

The momentum factor, $MOM,$ is calculated using six value-weighted portfolio which were creased based on size and prior returns that are listed in the NYSE, AMEX, and NASDAQ. Than to find the MOM, the two lowest prior average return portfolio is subtracted from the two highest prior average return portfolios. This factor measure if the fund experiences momentum, which is if the price continues experience the same trend as the previous periods.

Data Envelopment Analysis

While these traditional ratios and methodology are useful, another method is used to accurately measure performance of mutual fund and take account of the ethical aspect that our funds possess. Thus the performance that is used here and which have been used in other mutual fund related literature (such as Basso and Funari 2001, 2003; Rubio et al. 2014) is the Data Envelopment Analysis (DEA) approach. The DEA model is beneficial because it can use multiple inputs and outputs to measure the relative performance of the decision making unit. The DEA is a popular operational management methodology that tests the efficiency of decision making. The premise of this methodology is to compare a producer's (or a decision making unit, DMU) efficiency with the "best" producer (or the efficient frontier). Thus the produces takes on a set of inputs to produce a set of outputs. An efficient decision maker thus would seek to maximize output while minimizing input. In the investment environment this would be similar to maximizing return while minimizing risk. While there is a number of way to formulate the DEA the most direct formulation will be given: were X_i is a vector of inputs that produces a vector of outputs, Y_i, where *i* is the number of funds. Therefore to measure the efficiency of DMU_0 , by estimating the performance, P, fund the following linear program would be used:

$$P = Min \quad \theta$$

$$s.t. \sum \lambda_i X_i \leq \theta\, X_0 \tag{6}$$

$$\sum \lambda_i Y_i \geq Y_0$$

$$\lambda \geq 0$$

where λ is the exogenous weight fitted to DMU *i* in its attempt to dominate DMU_0 which efficiency is represented by θ. Thus when P=1, then they have reach efficiency level.

In this DEA model, input will represent a risk measure while output will represent all return measurements. The literature proposes 3 different feasible input for risk: standard deviation, lower partial moments (LPM), and maximum drawdown periods (MDP). The proposal for outputs are: expected returns, the upper partial moments (UPM), and the maximum consecutive gain (MCG). LPM demonstrates the risk of holding an investment security while UPMs captures the gains of holding the investment. This study utilizes partial movements to differentiate between inputs and outputs. In order to estimate both the lower and upper partial moments as the m^{th} root of these variables, the mean return, r_{min}, is employed to distinguish between the upside and downside of investment.

$$LPM_{j.m} = \frac{1}{T}\sum_{t=1}^{\tilde{T}}(r_{min} - \underline{r}_{t,j})^m$$

$$UPM_{j.m} = \frac{1}{T}\sum_{t=1}^{\hat{T}}(\bar{r}_{t,j} - r_{min})^m \quad \forall \; m = 0, \dots, 4 \tag{7}$$

Where r_{min} is the target rate, $\underline{r}_{t,j}$ is the monthly return of fund j that is below the target rate, $\bar{r}_{t,j}$ is the monthly return of fund j that is above the target rate. Utilizing partial movements, BBC is found to determine how efficiently funds take risk to produce return. These scores are than used to compare with the average of their own fund category.

MDP and MCG can also be used if Net Present Value (NPV) of the funds are available. Funds j's MDP represents the maximum number of consecutive months when the fund's net asset value is lower than the historic high. Conversely, MCG is when the maximum number of consecutive months are above the minimum target rate.

Using old traditional methods and a newer efficient method we can study the performance of SRI funds as well as Christian mutual funds.

5. Results

Sharpe Ratio

Table 3 shows the Sharpe ratio when the differential benchmark is either the risk-free rate or the S&P 500 index. Social Responsible Mutual funds has a positive Sharpe ratio when comparing with risk free rate which is statistically significant in the 1% level. SRI funds has an average excess return of 4.6% per unit of risk. Catholic mutual funds' Sharpe ratio is significant in the 10% level. Catholic funds has an average excess return of 4.8% per unit of risk. The Sharpe Ratio calculated for the comprehensive Christian mutual fund data set and that of the Protestant mutual funds are statistically insignificant. Thus while Social Responsible funds and Catholic funds returns are more volatile, they tend to be compensated for the extra risk they bare.

The Sharpe ratio for all funds using the S&P 500 index as the benchmark are all negatively significant in the 1% level. Thus passive investing and simply following the S&P 500 index seems to be a better investment choice. During the Financial Crisis, however, the Sharpe ratio calculation tells a different story. In 2008, the Sharpe ratio for all four funds are positively significant in the 1% level compared to the S&P 500 Index. The comprehensive Christian mutual fund gets compensated with almost 16% excess return per unit of risk, while the Social Responsible funds has a 12% excess return per unit of risk. Thus even though SRI funds held more risk they did not get compensated as well as the Christian mutual funds. However, SRI funds were the only funds that continued to receive compensation in 2009. Thus, in 2008 at the beginning of the Financial Crisis, the screened funds performed better than the market, yet it would have been a better investment choice to go with a risk-free asset. In 2009, after the peak of the Financial Crisis and due to the government borrowing and intervention, the funds became a better investment opportunity than the risk free market. In 2014, these two types of screened mutual funds are not deemed as good investment for the risk they take compared to both the risk-free asset and the S&P.

Table 3: Sharpe Ratio

The Sharpe ratio for Christian Mutual funds and Social Responsible mutual funds is used to compare the performance of mutual funds to the risk-free asset, one-month Treasury bill rate, and the market, S&P 500 Index.

$$Sharpe\ Ratio = (R_p - R_f)/\sigma$$

The Sharpe ratio calculates the excess return per unit of risk. This comparison is done throughout the entire sample as well as by year. The t-statistics has also been calculated.

	Christian Mutual Fund						*Social Responsible Mutual Funds*			
	Rf	**T-stat**	**S&P**	**T-stat**			**Rf**	**T-stat**	**S&P**	**T-stat**
CMF	0.0143	1.455372	-0.147	-14.9608		***SRI***	0.04564	5.531485	-0.1052	-12.75
2005	-0.0296	-0.76331	-0.0558	-1.43895		2005	-0.0128	-0.38379	-0.0653	-1.95791
2006	-0.0077	-0.20647	-0.2885	-7.73589		2006	0.113	3.591196	-0.2146	-6.82009
2007	-0.1548	-4.61813	-0.1452	-4.33173		2007	-0.0826	-2.8482	-0.0671	-2.31373
2008	-0.544	-16.864	0.1593	4.9383		2008	-0.4996	-18.3361	0.1214	4.455562
2009	0.3401	10.65767	-0.0105	-0.32904		2009	0.3751	14.41087	0.0654	2.512586
2010	0.2304	7.56821	-0.0387	-1.27122		2010	0.188	7.543463	-0.0663	-2.66027
2011	-0.0298	-1.0044	-0.0875	-2.94915		2011	-0.0866	-3.52516	-0.1771	-7.20908
2012	0.234	8.153146	-0.1683	-5.86399		2012	0.2676	11.04964	-0.0929	-3.83599
2013	0.2684	9.56122	-0.1989	-7.08542		2013	0.4158	17.67517	-0.3275	-13.9216
2014	-0.0585	-2.13505	-0.4158	-15.1753		2014	0.0044	0.188534	-0.3419	-14.6499
2015	-0.5379	-5.66713	0.3376	3.556837		2015	-0.6893	-8.52617	0.2383	2.947608

	Protestant Mutual Funds						*Catholic Mutual Funds*			
	Rf	**T-stat**	**S&P**	**T-stat**			**Rf**	**T-stat**	**S&P**	**T-stat**
Prot	*0.0085*	*0.7995*	*-0.1462*	*-13.7514*		***Cath***	*0.0476*	*1.85029*	*-0.1535*	*-5.96679*
2005	-0.035	-0.84726	-0.0596	-1.44276		2005	0.00866	0.076972	-0.0201	-0.17865
2006	-0.0273	-0.68414	-0.305	-7.64328		2006	0.134	1.278279	-0.1618	-1.54347
2007	-0.1535	-4.29526	-0.1409	-3.94268		2007	-0.1681	-1.73884	-0.1959	-2.02641
2008	-0.5581	-16.2137	0.1382	4.014942		2008	-0.4479	-4.84478	0.3863	4.178473
2009	0.3337	9.786005	-0.0225	-0.65983		2009	0.3827	4.22706	0.1114	1.230453
2010	0.221	6.706904	-0.0475	-1.44153		2010	0.2824	3.549713	0.024	0.301675
2011	-0.0382	-1.18112	-0.0992	-3.06719		2011	0.0144	0.193196	-0.0121	-0.16234
2012	0.245	7.77465	-0.1668	-5.29311		2012	0.1855	2.66888	-0.1752	-2.52069
2013	0.2656	8.61871	-0.4967	-16.1179		2013	0.2811	4.131309	-0.5105	-7.50279
2014	-0.0461	-1.54004	-0.4096	-13.6833		2014	-0.1131	-1.66222	-0.4481	-6.5857
2015	-0.4263	-4.11109	0.553	5.332939		2015	-1.1746	-4.98341	-0.4117	-1.7467

Jensen's Alpha

Table 4 demonstrates the result for Jensen's Alpha. SRI funds underperforms the market by .3868 basis points, which is approximately a -4.64% underperformance after adjusting for systematic risk. Christian mutual funds also underperforms, but by .46 basis points, this is approximately a -5.52% underperformance. In fact, all four funds have a statistically negative alpha and a beta less than 1. Thus although these funds had lower risk than the market's beta, they still underperformed the market. Of the four funds, SRI funds was slightly less negative. In fact SRI is statistically significantly higher than Christian mutual funds in 5% level. Protestant and Catholic funds have similar returns, with Catholic funds having a higher market risk however the difference is not statistically significant. This model has decent r-square ranging from 57%-68%, thus does a fairly good job in predicting the variation of the returns using just these two variables. Panel B presents the fund's alpha per year. Here we find interesting results during the Great Recession. SRI funds performed similar to the market in 2008, but outperform the market slightly in 2009 with a .38 basis point increase per month, or a 4.6% annual increase. While the comprehensive Christian mutual fund did not outperform the market any year, Catholic funds outperform the market in 2008 with an excessive .69 basis points per month and performed as well as the market in 2009 and 2010. In 2008, they definitely outperform the Protestant fund However, in the overall data, Catholics had the worst alpha and the highest beta. Thus, Catholic funds tend to take more risk, and while it paid off during the financial crisis, in recent years this has not been the case. In 2014, they statistically significantly (in 5% level) performed worse than Protestant funds.

Table 4: Jensen's Alpha

Jensen's alpha measure the abnormal return of the portfolio.

Jensen's alpha = $\propto = R_p - [R_f + \beta(R_M - R_f)]$

Where R_p is the portfolio return, R_f is the risk-free rate, and R_M is the return of the market. The alpha coefficient represents the abnormal performance of the fund, while beta of the excess return is systematic risk. Panel A runs the regression for the entire data set, while Panel B runs the data for each individual year.

**** 1% Statistically Significant*
*** 5% Statistically Significant*
** 10% Statistically Significant*

Panel A: For Entire Data

	SRI	CMF	Protestant	Catholic
Alpha	-0.3868***	-0.4605***	-0.4619***	-0.4661***
Rm-Rf	0.8394***	0.7595***	0.7452***	0.8482***
R-square	63.58%	58.77%	57.30%	67.96%

Panel B: Alpha per Year

	SRI	CMF	Protestant	Catholic
2005	-0.3138***	-0.2930***	-0.2956***	-0.2821
2006	-0.4605***	-0.7422***	-0.7750***	-0.5223**
2007	-0.3316***	-0.5846***	-0.5857***	-0.5784***
2008	0.0572	-0.2486*	-0.3809***	0.6939**
2009	0.3846***	0.1736	0.1691	0.2014
2010	-0.2309***	0.0378	0.0231	0.1142
2011	-0.5347***	-0.2546***	-0.2922***	-0.0531
2012	-0.2410***	-0.2752***	-0.2417***	-0.4248***
2013	-0.6199***	-0.9992***	-0.9578***	-1.2000***
2014	-0.8241***	-0.9870***	-0.9099***	-1.3851***

Fama-French and Carhart

The Fama-French three factor model, as seen in table 5, and the Carhart model, table 6, are performed for robustness. The r-square did slightly increase although not as much as

Table 5: Fama and French Three Factors

The Fama-French three factor model captures performance of the fund through alpha and systematic risk, but also looks at additional risk involved with size and value.

$$\left(R_p - R_f\right) = \propto\ + \beta_{i,1,t}\left(R_{M,t} - R_{f,t}\right) + \ \beta_{S,i,}SMB_t + \ \beta_{H,i}HML_t$$

where (R_p- R_f) is the excess return of the portfolio , $\left(R_{M,t} - R_{f,t}\right)$ is the excess return of the market, SMB_tvariable is Fama and French's "small minus big" factor, HML is the "high minus low" factor. The regression is run on the whole sample as well as individual years for the Social Responsible funds (SRI), Christian mutual funds (CMF), and the Protestant and Catholic sects. The bottom reflects r-square.

**** 1% Statistically Significant*
*** 5% Statistically Significant*
** 10% Statistically Significant*

Panel A: Entire Sample

	SRI	CMF	Protestant	Catholic
Alpha	-0.4031***	-0.4708***	-0.4748***	-0.4518***
Rm-Rf	0.8534***	0.7744***	0.7676***	0.8163***
SMB	0.0220**	-0.0139	-0.0351**	0.1146***
HML	-0.1038***	-0.0637***	-0.0749***	0.0155
R-square	63.82%	58.88.%	57.49%	68.26%

Panel B: Alpha per Year

	SRI	CMF	Protestant	Catholic
2005	-0.2730***	-0.3282***	-0.3324***	-0.3083
2006	-0.2151**	-0.2994**	-0.3090**	-0.2402
2007	-0.6867***	-1.0903***	-1.1694***	-0.5357*
2008	0.0983	-0.0313	-0.0998	0.4623
2009	0.2564**	0.1982	0.2199	0.0377
2010	-0.2180***	-0.003	0.0076	-0.0425
2011	-0.6965***	-0.3775***	-0.4402***	-0.0407
2012	-0.2398***	-0.2224***	-0.1746**	-0.4383***
2013	-0.7390***	-1.2780***	-1.2284***	-1.5190***
2014	-1.0374***	-1.2452***	-1.1651***	-1.6287***

conveyed by Fama and French. When the Fama and French size and value factors are included, the underperformance observed in the Jensen's alpha is only aggravated more. SRI (Christian) funds underperform the market by .43 (.47) basis points per month, -4.8% (.65%) annually. Again, the alphas are all negative and significant while the betas are all less than one and also significant when looking at the whole sample date. This result is similar to mutual fund performance against the market. In 2008, the alphas cease to be significant, the market betas are significant and still lower than one, though they are higher than other years. So while they still had less market risk than the market portfolio, their funds had more market risk than previous years. In 2009, only the SRI funds intercept coefficient was significant in the 5% level. As in the Jensen's alpha case, the alpha is positively significant thus they outperform the market by .2564 basis points per month. This seems to support early literature that mutual funds perform better than the market during economic downturns. However this was short lived since in 2010, the alpha again became negatively significant. During the financial crisis, the alpha for the remaining funds remained insignificant. Thus from 2008-2010, Christian funds performed no better or worse than the market.

While both SRI and CMF have negative alphas, in most cases CMF exhibited an inferior alpha (statistically significant in 10% level). In fact, SRI has a superior performance to the separate Christian funds as well. This pattern remain in each individual year. This seems to indicate that the extra pressure of religious screening hurts the performance of religious funds. When dividing the Christian funds between Protestant and Catholic, at first glance, it seems that Catholics have performed better than Protestants even though they tend to involve more risk. However, when studying each individual year, the difference of the coefficients are only statistically significant (5% level) in 2007. As indicated by the Jensen's alpha, in 2014 Protestant funds perform better than Catholic funds (statistically significant in 10% level). In other words, since 2012 Catholics have had the more negative alpha while still having the larger beta. Thus the Catholic funds have taken on more risk and it has only made their performance worse.

The coefficient for size premium are positive for SRI and Catholic funds. These funds had higher risk with a positive SMB coefficient indicates that they invest in small market equity portfolio. The slightly more conservative Protestant funds has a negative SMB coefficient. Thus as firm size increases, the SMB coefficient decreases. The SMB value would decrease and thus lower the funds return. Historically, the SMB factors have received a size premium of about 3.3%, yet in recent years due to financial turmoil this size premium has significantly decrease. In 2008, at the height of the financial crisis, the SMB coefficient for SRI and Catholic funds have become insignificant, while the CMF and Protestant SMB funds have become more negative. Protestant fund tend to gravitate more to large cap firms during crisis since they tend to perform better. The HML coefficient for SRI and Christian mutual funds are also negative and significant. Thus as equity firms have higher book-to-market value, it will have a lower coefficient. This seems to indicate that growth stocks outperform value stocks during the extant of this period.

Table 6 demonstrates the result of Carhart four factor model. Again, the r-square is only slightly higher than before. Thus even for controlling for momentum the alphas are negatively significant. The momentum factor is negatively significant in all funds signifying an absence of the momentum effect articulated in other research. This seems to indicate mean reversion. All the alphas are negatively significant for the whole sample period. Christian mutual funds perform worse than SRI funds, and Catholic funds perform better than Protestant funds. During the Financial Crisis alphas become insignificant. Even though alphas are no longer significant with

Table 6: Carhart Four Factor Model

Carhart is an extension of the Fama-French three-factor model to include the momentum factor.

$$(R_p - R_f) = \propto\ + \beta_{i,1,t}(R_{M,t} - R_{f,t}) + \beta_{S,i,}SMB_t + \beta_{H,i}HML_t + \beta_{M,i}MOM_t$$

where (R_p- R_f) is the excess return of the portfolio , $(R_{M,t} - R_{f,t})$ is the excess return of the market, SMB_tvariable is Fama and French's "small minus big" factor, HML is the "high minus low" factor, and MOM is the momentum factor. The regression is run on the whole sample as well as individual years for the Social Responsible funds (SRI), Christian mutual funds (CMF), and the Protestant and Catholic sects. The bottom reflects r-square.

**** 1% Statistically Significant*
*** 5% Statistically Significant*
** 10% Statistically Significant*

Panel A: Entire Sample

	SRI	CMF	Protestant	Catholic
Alpha	-0.3894***	-0.4543***	-0.4597***	-0.4240***
Rm-Rf	0.8353***	0.7557***	0.7494***	0.7929***
SMB	0.0276**	-0.0085	-0.0295**	0.1204***
HML	-0.1383***	-0.1029***	-0.1148***	-0.0233
MOM	-0.0654***	-0.0667***	-0.0632***	-0.0950***
R-squared	64.19%	59.31%	57.88%	68.98%

Panel B: Alpha per Year

	SRI	CMF	Protestant	Catholic
2005	-0.2680***	-0.3923***	-0.4011***	-0.3448
2006	-0.4209***	-0.7474***	-0.8302***	-0.187
2007	-0.3537***	-0.4578***	-0.5172***	-0.0421
2008	0.0652	-0.0786	-0.1523	0.4733
2009	0.0063	0.0369	0.072	-0.2189
2010	-0.2117***	0.0141	0.0248	-0.0223
2011	-0.8746***	-0.3981***	-0.4496***	-0.1203
2012	0.0856	0.1459*	0.2090**	-0.1409
2013	-0.6459***	-1.1919***	-1.1348***	-1.4665***
2014	-0.9941***	-1.1880***	-1.1126***	-1.5781***

the extra momentum factor, the evidence point out that they at least perform similarly to the market. The r-square is also the greatest during these times. The same performance trend is shown in recent years. While SRI continues to outperform the religious funds, Catholic have perform worse than Protestant funds.

Data Envelopment Analysis

The DEA was performed based on three years' worth of data which made some funds invalid to find the estimated monthly BCC scores. These scores are than averaged out and compared the group average. The objective of this analysis is to measure the relative efficiency of each fund compared to other funds in the same group. Table 7 depicts the number of funds that are efficient. As can be seen from the result, Catholic funds are the most efficient. About half of the Catholic funds are efficient. Of course this result may be skewed by the lack of number of valid funds. Comparing Social responsible funds and Christian mutual funds, Social responsible funds are exceedingly more efficient than Christian mutual funds. About 44% of SRI funds are efficient compared to 37% of Christian mutual funds. However, this still indicates that the majority of funds have shown an inefficient BCC score (62 of 140 funds). This is consistent with finding previous findings. Screening funds have hurt performance and efficiency. Furthermore, while Christian mutual funds, as a group, are less efficient than SRI funds this is mostly contributed to the larger number of protestant funds. Catholic funds, who tend to perform better than their Protestant counterpart, are also seen to be more efficient.

There are a number of funds that have 1 as their BBC score for several months. The Protestant fund that has the most months with a BBC score of fund is the Guidestone Low Duration bond funds. In fact, many of the higher BBC rated funds are conservative in nature either focusing on bonds, income funds or conservative allocation type funds. While the Catholic funds have higher proportion of efficiency, the number of funds are less along with the number of months. The bond which has a BBC score of 1 for most months is the Epiphany FF Strategic Income. However, true to the Catholic's fund characteristic, the other funds are not as conservative as that of Protestant, the efficient funds are small cap, value funds and growth funds. SRI funds who has a measure of 1 BBC score also lean towards conservative investment. The fund which scored the most 1 as there BBC score is the Access Capital Community Investment which is an intermediate Government bond fund. Thus the DEA method tells us that while the Catholic have a greater proportion of efficient funds, mainly efficient funds tend to lean toward conservative investment.

Table 7: Data Envelopment Analysis

The Data Envelopment Analysis (DEA) measure the efficiency of fund's performance by comparing the funds inputs and output. To measure the efficiency of the decision making unit they try to maximize return by minimizing risk. This study utilizes partial movements to differentiate between inputs and outputs to calculate BBC Score. In order to calculate BBC, funds need to have at least 3 years of information. The average of efficiency scores are found for each fund type and below shows the percentage of funds that are above the average efficiency score.

Fund Type	Number of Valid Funds	Above Average	Percentage
CATH	15	8	53.33%
CMF	98	36	36.73%
PROT	83	33	39.76%
SRI	140	62	44.29%

6. Conclusion

While religion was the first to screen their investment, they have now been swallowed up into a general category along with other "social responsible" investing. However, while what was deemed socially responsible by the public at one time was seen synonymous to Christian values, this is obviously not the case now. While SRI funds invest heavily in environmentally friendly and social awareness funds, Christians have taken another path to focus on investing based on biblical principles. Christian funds are more concerned with staying away from moral issues such as abortion, pornography, and emphasizes on staying true to biblical truth or follow guidelines submitted by higher authority of the church. Therefore, I argue that just as Islamic funds are being seen as separate from SRI grouping, so should Christian funds.

According to the DEA method, these screened funds are not highly efficient. In most cases, less than half of the fund exceed the average mark of its group. Furthermore, SRI fund and Christian funds tend to perform worse than the market during normal economic times. However, during recession screened mutual funds performed as well and in times better than the market. This follows literature which determines that mutual funds are attractive as an insurance mechanism during market downturns. Christian mutual funds performed worse than SRI funds. This may suggest that their lower NAV and their exclusionary screening in "sin stocks" and moral ambiguous stocks neutralizes diversification benefits and results in lower returns than even SRI funds.

This research was limited by data provided and results could be improved with a more detail and comprehensive dataset. Yet the results have proven interesting. We have showed that moral screening has an effect on result. As literature suggest, Catholics tend to take more risk than Protestants and in many cases taking extra risk paid off. In fact, Catholic funds are the most efficient subset funds. In many cases, Christian organizations invest in Christian funds in their retirement fund. This information should be useful to them. While they may be abstain from sinful investing and maintain their moral principles, this in fact may hurt them financially.

7. Reference

Adhikari, B., & Agrawal, A. (2014). Does Local Religiosity Matter for Bank Risk-Taking? *Available at SSRN 2465052*.

"America's Largest Family of Catholic Mutual Funds." *Home: Ave Maria Mutual Funds*. Schwartz Investment Councel Inc., 2014. Web. 21 Aug. 2014.

Arruñada, Benito. (2010). "Protestants and Catholics: Similar Work Ethic, Different Social Ethic*." *The Economic Journal* 120.547 890-918.

Barnett, Michael L., and Robert M. Salomon. (2006). "Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance." *Strategic Management Journal* 27.11 1101-1122.

Barro, R. J., & McCleary, R. (2003). *Religion and economic growth* (No. w9682). National Bureau of Economic Research.

Basso, A., & Funari, S. (2001). A data envelopment analysis approach to measure the mutual fund performance. *European Journal of Operational Research*, *135*(3), 477-492.

Basso, A., & Funari, S. (2003). Measuring the performance of ethical mutual funds: a DEA approach. *Journal of the operational research society*, *54*(5), 521-531.

Becker, Sascha O., Ludger Woessmann (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. Quarterly Journal of Economics124 (2): 531-596.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance, 52*(1), 57-82.

Cochran, Philip L., and Robert A. Wood. (1984). "Corporate social responsibility and financial performance." *Academy of management Journal* 27.1 42-56.

DiBartolomeo, Dan, and Lloyd Kurtz. (1999). "Managing risk exposures of socially screened portfolios." *Northfield Information Services* 1-17.

David Diltz, J. (1995). "The private cost of socially responsible investing." *Applied Financial Economics* 5.2 69-77.

Elfakhani, Said, and M. Kabir Hassan. "Performance of Islamic mutual funds." *Economic Research Forum, 12th Annual Conference*. 2005.

Fama, Eugene F., and Kenneth R. French. (1993). "Common risk factors in the returns on stocks and bonds." *Journal of financial economics* 33.1 3-56.

Fama, Eugene F., and Kenneth French. (2008). "Mutual fund performance." *Journal of Finance* 63 389-416.

Forte, Gianfranco, and Federica Miglietta. (2007). "Islamic mutual funds as faith-based funds in a socially responsible context."

Friedman, Howard Steven. "5 Religions With The Most Followers." *The Huffington Post*. TheHuffingtonPost.com, 25 Apr. 2011. Web. 14 May 2014.

Friedman, Milton. (1973)."The social responsibility of business is to increase its profits." *New York*

Geczy, Christopher C., Robert F. Stambaugh, and David Levin. (2003). *Investing in socially responsible mutual funds*. Rodney L. White Center for Financial Research, The Wharton School, University of Pennsylvania.

Goldreyer, Elizabeth F., and J. David Diltz. (1999). "The performance of socially responsible mutual funds: incorporating sociopolitical information in portfolio selection." *Managerial Finance* 25.1 23-36.

Greening, Daniel W., and Daniel B. Turban. (2000)."Corporate social performance as a competitive advantage in attracting a quality workforce." *Business & Society* 254-280.

Guerard Jr, John B. (1997). "Is there a cost to being socially responsible in investing?." *The Journal of Investing* 6.2 11-18.

Hamilton, Sally, Hoje Jo, and Meir Statman. (1993). "Doing well while doing good? The investment performance of socially responsible mutual funds." *Financial Analysts Journal* 62-66.

Hilary, Gilles, and Kai Wai Hui. (2009). "Does religion matter in corporate decision making in America?." *Journal of Financial Economics* 93.3 455-473.

Hoepner, Andreas GF, Hussain G. Rammal, and Michael Rezec. (2011). "Islamic mutual funds' financial performance and international investment style: evidence from 20 countries." *The European Journal of Finance* 17.9-10 829-850.

Iannaccone, L. R. (1998). Introduction to the Economics of Religion. *Journal of economic literature*, 1465-1495.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65-91.

Jensen, Michael C. (1968). "The performance of mutual funds in the period 1945–1964." *The Journal of finance* 23.2 389-416.

Kumar, A., Page, J. K., & Spalt, O. G. (2011). Religious beliefs, gambling attitudes, and financial market outcomes. *Journal of Financial Economics*, *102*(3), 671-708.

Lipford, J. W., & Tollison, R. D. (2003). Religious participation and income. *Journal of Economic Behavior & Organization*, *51*(2), 249-260.

"LKCM Aquinas Funds." *LKCM Aquinas Funds*. Quasar Distributors, LLC., 2014. Web.

McCleary, R. M. (2008). Religion and economic development. *Policy Review*, *148*, 45-57.

McGuire, S. T., Omer, T. C., & Sharp, N. Y. (2011). The impact of religion on financial reporting irregularities. *The Accounting Review*, *87*(2), 645-673.

McWilliams, Abagail, and Donald Siegel. (1997). "The Role of Money Managers in Assessing Corporate Social Reponsibility Research." *The Journal of Investing* 98-107.

"New Covenant Funds." *New Covenant Funds*. Presbyterian Foundation Group. Web.

"New Research Finds That Protestants And Catholics Have A Similar Work Ethic." *Royal Economic Society*. Wiley-Blackwell, Sept. 2010. Web. 20 July 2014

Newport, Frank. "In U.S., Four in 10 Report Attending Church in Last Week." *Gallup Politics*. Gallup Inc., 24 Dec. 2013. Web.

Peifer, Jared L. (2011). "Morality in the Financial Market? A Look at Religiously Affiliated Mutual Funds in the USA." Socio-Economic Review 9(2).

"Performance + Values." *GuideStone Funds*. Foreside Funds Distributors LLC, 2014. Web.

Rouwenhorst, K. G. (2004). The origins of mutual funds. Yale ICF Working Paper No. 07-48

Rubio, Jose F., Neal C. Maroney, and M. Kabir Hassan. *Can Efficiency of Returns Be Considered as a Pricing Factor?* Diss. U of New Orleans, 2014. N.p.: n.p., n.d. Print.

Rubio, Jose Francisco, M. Kabir Hassan, and Hesham Jamil Merdad. (2012). "Non-parametric performance measurement of international and Islamic mutual funds." *Accounting Research Journal* 25.3 208-226.

Sharpe, William F. (1966). "Mutual fund performance." *Journal of business* 119-138.

Sharpe, W. F. (1994). The Sharpe Ratio. *The journal of portfolio management*, *21*(1), 49-58.

Shu, Tao, Johan Sulaeman, and P. Eric Yeung. (2012). "Local religious beliefs and mutual fund risk-taking behaviors." *Management Science* 58.10 1779-1796.

Statman, Meir. (2000). "Socially responsible mutual funds." *Financial Analysts Journal* 30-39.

"Thrivent Financial | Annuities, Insurance, Mutual Funds, IRAs & More." *Thrivent Financial for Lutherans*. 10 July 2014. Web.

"Timothy Plan® | Biblically Responsible Investing." *Timothy Plan® | Biblically Responsible Investing*. 2014. Web.

Treynor, Jack L. (1965). "How to rate management of investment funds." *Harvard business review* 43.1 63-75.

Waddock, Sandra A., and Samuel B. Graves. (1997). "The corporate social performance." *Strategic management journal* 8.4 303-319.

Weber, Max. (1904). "1930." *The protestant ethic and the spirit of capitalism.*

Vita

The author was born in New Orleans, Louisiana in July, 1988. She obtained her Bachelor of Science degree in Finance and International Business from LeTourneau University in August, 2009. She joined the University of New Orleans Financial Economics Program to pursue a PhD in Fall, 2011. She earned her Master of Science in Financial Economics in May 2013 and became a Teaching Assistant thereafter. She was hired tenure-track Assistant Professor in Fall, 2015 in the University of Houston Downtown. She obtained her PhD in Financial Economics from the University of New Orleans in December, 2015.