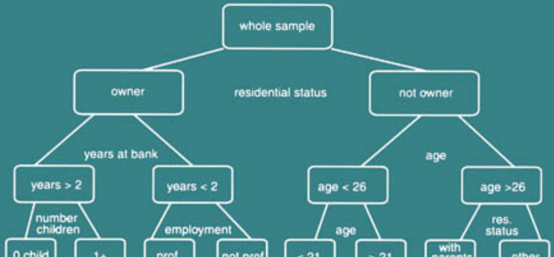
Credit Scoring and Its Applications

Lyn C. Thomas
David B. Edelman
Jonathan N. Crook



siam
Monographs on Mathematical Modeling and Computation

Credit Scoring and Its Applications

SIAM Monographs on Mathematical Modeling and Computation

About the Series

In 1997, SIAM began a new series on mathematical modeling and computation. Books in the series develop a focused topic from its genesis to the current state of the art; these books

- present modern mathematical developments with direct applications in science and engineering;
- describe mathematical issues arising in modern applications;
- develop mathematical models of topical physical, chemical, or biological systems;
- present new and efficient computational tools and techniques that have direct applications in science and engineering; and
- illustrate the continuing, integrated roles of mathematical, scientific, and computational investigation.

Although sophisticated ideas are presented, the writing style is popular rather than formal. Texts are intended to be read by audiences with little more than a bachelor's degree in mathematics or engineering. Thus, they are suitable for use in graduate mathematics, science, and engineering courses.

By design, the material is multidisciplinary. As such, we hope to foster cooperation and collaboration between mathematicians, computer scientists, engineers, and scientists. This is a difficult task because different terminology is used for the same concept in different disciplines. Nevertheless, we believe we have been successful and hope that you enjoy the texts in the series.

Mark Holmes

Johnny T. Ottesen, Mette S. Olufsen, and Jesper K. Larsen, *Applied Mathematical Models in Human Physiology*

Ingemar Kaj, *Stochastic Modeling in Broadband Communications Systems*

Peter Salamon, Paolo Sibani, and Richard Frost, *Facts, Conjectures, and Improvements for Simulated Annealing*

Lyn C. Thomas, David B. Edelman, and Jonathan N. Crook, *Credit Scoring and Its Applications*

Frank Natterer and Frank Wübbeling, *Mathematical Methods in Image Reconstruction*

Per Christian Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*

Michael Griebel, Thomas Dornseifer, and Tilman Neunhoeffer, *Numerical Simulation in Fluid Dynamics: A Practical Introduction*

Khosrow Chadan, David Colton, Lassi Päivärinta, and William Rundell, *An Introduction to Inverse Scattering and Inverse Spectral Problems*

Charles K. Chui, *Wavelets: A Mathematical Tool for Signal Analysis*

Credit Scoring and Its Applications

Lyn C. Thomas
University of Southampton
Southampton, United Kingdom

David B. Edelman
Royal Bank of Scotland
Glasgow, United Kingdom

Jonathan N. Crook
University of Edinburgh
Edinburgh, United Kingdom

siam
Society for Industrial and Applied Mathematics
Philadelphia

10 9 8 7 6 5 4 3 2

Library of Congress Cataloging-in-Publication Data
Thomas, L. C.
Credit Scoring and its Applications / Lyn C. Thomas, David B. Edelman, Jonathan N. Crook.
p. cm. - (SIAM monographs on mathematical modeling and computation)
Includes bibliographical refreences and index.
ISBN 0-89871-483-4 (pbk.)
1. Credit scoring systems. I. Edelman, David B. II. Crook, Jonathan N. III. Title.
IV. Series

HG3751.5 .T47 2002
658.8'8–dc21

2001049846

Disclaimer: This eBook does not include ancillary media that was packaged with the printed version of the book.

To Margery, Matthew, Beth, and Stephen

To Kate, Rachael, and Becky, who taught me which decisions really matter

To Monica, Claudia, and Zoë, who provide the light in my life, and to Susan, who keeps that light shining brightly

This page intentionally left blank

Contents

This page intentionally left blank

Preface

Credit scoring is one of the most successful applications of statistical and operations research modeling in finance and banking, and the number of scoring analysts in the industry is constantly increasing. Yet because credit scoring does not have the same glamour as the pricing of exotic financial derivatives or portfolio analysis, the literature on the subject is very limited. However, credit scoring has been vital in allowing the phenomenal growth in consumer credit over the last 40 years. Without an accurate and automatically operated risk assessment tool, lenders of consumer credit could not have expanded their loan books in the way they have.

Credit scoring was one of the earliest financial risk management tools developed. Its use by U.S. retailers and mail-order firms in the 1950s is contemporary with the early applications of portfolio analysis to manage and diversify the risk inherent in investment portfolios. Also, credit scoring could claim to be the grandfather of data mining because it was one of the earliest uses of data on consumer behavior. In fact, the commonest techniques used in data mining—segmentation, propensity modeling, and clustering—are also techniques that have been used with considerable success in credit scoring.

This book outlines the basic theory underpinning the development of credit-scoring systems. It also looks at the practical issues that arise in the building and monitoring of such systems. The areas where scoring is used and the possible developments in scoring usage are also discussed. It is hoped that such an introduction will be of use to those studying scoring as part of courses in operations research and statistics and to practitioners who might find it useful as a reference book. The need for such a book was brought home to us in the series of international conferences on credit scoring and credit control run by the Credit Research Centre at the University of Edinburgh over the past decade. At these conferences, there was a plethora of new uses and new ideas for scoring but a dearth of literature where these ideas were expounded.

The book developed out of a course in credit scoring run jointly for master's degrees in operations research and financial mathematics. The book could also be used as part of applications modules in degree and master's courses in statistics and operations research. In that case, one might concentrate on the chapters that outline the basic problems and the techniques used to develop risk assessment systems: Chapters 1, 2, 4–8, and 12. Those who want to look only at current areas of application and avoid the modeling aspects should look at Chapters 1, 2, 9–11, 13, and 14, while Chapter 3 gives an introduction to the economic theory underlying models of consumer indebtedness.

Thus the book as a whole should give practitioners a background in which to place the everyday decisions in which they will be involved when they commission, build, operate, and monitor credit and behavioral scoring systems. Because decisions must be made about what to include and exclude in such a text, we concentrate on consumer credit and behavioral scoring and do not investigate in great detail the risk-assessment tools used for company and trade credit, although there are some strong overlaps. Similarly, in the last few years, it has become clear that there are some interesting connections in the ways of dealing with credit risk in the pricing of bonds and other financial instruments and the techniques that are used or are being developed for consumer credit scoring. Again, we decided not to investigate such connections in this book, nor do we go deep into the relationship between credit scoring and the securitization of the loans that have been scored. We tried to focus primarily on assessing the risk involved in lending to consumers.

One way to understand the ideas in the text is to try them out on real data. For practitioners this is not a problem, but for those who do not have access to data sets, we have included a small data set.

Credit scoring is a fascinating subject. It affects almost everyone—in the U.S. or Europe, a typical person will be credit scored on average at least twice a month. Yet it is also an area where many techniques in statistics and operations research can be used. Finally, it is an area in which even a small improvement in performance can mean a tremendous increase in profit to the lender because of the volume of lending made by using scoring. Relevant, ubiquitous, profitable—credit scoring is all of these, and we hope this book convinces the reader that it is also interesting and thought provoking.

In writing this book we benefited considerably from comments and suggestions from many people. We especially thank Phil Bowers, David Hand, Alan Lucas, and Bob Oliver for their helpful suggestions and their encouragement. All the errors are, of course, ours.

Chapter 1

The History and Philosophy of Credit Scoring

1.1 Introduction: What is credit scoring?

Credit scoring is the set of decision models and their underlying techniques that aid lenders in the granting of consumer credit. These techniques decide who will get credit, how much credit they should get, and what operational strategies will enhance the profitability of the borrowers to the lenders.

Credit-scoring techniques assess the risk in lending to a particular consumer. One sometimes hears it said that credit scoring assesses the creditworthiness of the consumer, but this is an unfortunate turn of phrase. Creditworthiness is not an attribute of individuals like height or weight or even income. It is an assessment by a lender of a borrower and reflects the circumstances of both and the lender's view of the likely future economic scenarios. Thus some lenders will assess an individual as creditworthy and others will not. One of the longer-term dangers of credit scoring is that this may cease to be the case, and there will be those who can get credit from all lenders and those who cannot. Describing someone as uncreditworthy causes offense. It is better for the lender to state the reality, which is that the proposition of lending to this consumer represents a risk that the lender is not willing to take.

A lender must make two types of decision—first, whether to grant credit to a new applicant and, second, how to deal with existing applicants, including whether to increase their credit limits. The techniques that aid the first decision are called credit scoring, and these are outlined in Chapters 4 and 5, while the techniques that assist the second type of decision are called behavioral scoring. The latter are described in detail in Chapter 6.

In both cases, whatever the techniques used, the vital point is that there is a very large sample of previous customers with their application details and subsequent credit history available. All the techniques use the sample to identify the connections between the characteristics of the consumers and how "good" or "bad" their subsequent history is. Many of the methods lead to a scorecard, where the characteristics are given a score and the total of these scores says whether the risk of a consumer being bad is too great to accept. Other techniques described in this book do not lead to such scorecards but instead indicate directly the chance that the consumer is good and so whether accepting the account is worthwhile. Although these approaches do not lead to a scorecard, they go by the name of credit and behavioral scoring.

Although most of this text, with the exception of Chapter 11, concentrates on the credit applications of scoring, scoring has been used in a number of other contexts in the last decade.

In particular, it is proving very useful for targeting customers in direct mailing and other marketing techniques. With the advent of data warehousing, many organizations, especially in the financial and retail sectors, find that they have the information needed to apply scoring techniques. Similarly, data mining, the other much-hyped advance in using information systems, has as one of its most successful application areas response classification, which is essentially scoring applied in a different context.

1.2 History of credit

As soon as humans started to communicate, it is certain that they started borrowing and repaying. One could speculate that a hunter would take the best club belonging to the group and then share the meat killed with that club with the rest of the group. The first recorded instance of credit comes from ancient Babylon. Lewis (1992) records that on a stone tablet dated to around 2000 BC is the inscription "Two shekels of silver have been borrowed by Mas-Schamach, the son of Adadrimeni, from the Sun priestess Amat-Schamach, the daughter of Warad-Enlil. He will pay the Sun-God's interest. At the time of the harvest he will pay back the sum and the interest upon it." Thus farmers were already dealing with their cash flow problems by borrowing at planting to pay back at harvest.

By the time of the Greek and Roman empires, banking and credit institutions were well advanced, although the idea of a credit card offering XVIII.IX% APR might have proved difficult to promote. The next thousand years, the "Dark Ages" of European history, saw little development in credit, but by the time of the Crusades in the thirteenth century, pawn shops had been developed. Initially, these were charities that charged no interest, but merchants quickly saw the possibilities, and by 1350 commercial pawn shops charging interest were found throughout Europe. These pawn shops, which lend on almost anything left on deposit, and their three-ball sign can still be found in most European and South American countries. During the Middle Ages there was an ongoing debate on the morality of charging interest on loans—a debate that continues today in Islamic countries. The outcome of the debate in Europe was that if the lender levied small charges, this was interest and was acceptable, but large charges were usury, which was bad. Even Shakespeare got into this debate with his portrait of the Merchant of Venice. Also at this time, kings and potentates began to have to borrow in order to finance their wars and other expenses. Lending at this level was more politics than business, and the downside for poor lending could be severe. The abbots who dealt with Henry VIII were given the choice, "Your monasteries or your life."

The rise of the middle classes in the 1800s led to the formation of a number of private banks, which were willing to give bank overdrafts to fund businesses and living expenses. However, this start of consumer credit was restricted to a very small proportion of the population. Mail-order and other enterprises began as clubs, mainly in Yorkshire. If an item costing £1 was desired by each of 20 people, then they would form a club and each pay in 5p (1 shilling in old money) a week. After one week they had enough to buy one of the items and they drew lots to determine whom should receive the item. This continued each week until week 20, when the last nonrecipient got the item. If they had all saved separately, then none of them would have had the item until week 20 without credit being extended. Therefore, in some sense this was a credit union with credit extended by ordinary people to other people, which allowed the lucky person to get his item when 19 further payments were due.

The real revolution started in the 1920s, when consumers started to buy motor cars. Here for the first time was an article that most consumers wanted and that was very mobile and so could not be deposited for security (like pawn shop lending) or even used as a security (like land and property, whose location the lender knows at all times). Finance companies were developed to respond to this need and experienced rapid growth before World War II. At the same time, mail-order companies began to grow as consumers in smaller towns demanded the clothes and household items that were available only in larger population centers. These were advertised in catalogues, and the companies were willing to send the goods on credit and allow customers to pay over an extended period.

Over the last half of the twentieth century, lending to consumers has exploded. Consumer credit has had one of the highest growth rates in any sector of business. The advent of credit cards in the 1960s was one of the most visible signs of this growth, and it is now difficult to function in society without one. In many cases purchases essentially can be made only if one uses a credit card—for example, over the Internet or on the telephone. It is important to remember, however, that credit cards account for less than 15% of consumer credit; far more is being borrowed through personal loans, hire purchases, overdrafts, and, of course, mortgages.

1.3 History of credit scoring

While the history of credit stretches back 5000 years, the history of credit scoring is only 50 years old. Credit scoring is essentially a way to identify different groups in a population when one cannot see the characteristic that defines the groups but only related ones. The first approach to solving this problem of identifying groups in a population was introduced in statistics by Fisher (1936). He sought to differentiate between two varieties of iris by measurements of the physical size of the plants and to differentiate the origins of skulls using their physical measurements. In 1941, Durand (1941) was the first to recognize that one could use the same techniques to discriminate between good and bad loans. His was a research project for the U.S. National Bureau of Economic Research and was not used for any predictive purpose.

During the 1930s, some mail-order companies had introduced numerical scoring systems to try to overcome the inconsistencies in credit decisions across credit analysts (Weingartner 1966, Smalley and Sturdivant 1973). With the start of the World War II, all the finance houses and mail-order firms began to experience difficulties with credit management. Credit analysts were being drafted into military service, and there was a severe shortage of people with this expertise. Hence the firms had the analysts write down the rules of thumb they used to decide to whom to give loans (Johnson 1992). Some of these were the numerical scoring systems already introduced; others were sets of conditions that needed to be satisfied. These rules were then used by nonexperts to help make credit decisions—one of the first examples of expert systems.

It did not take long after the war ended for some folks to connect the automation of credit decisions and the classification techniques being developed in statistics and to see the benefit of using statistically derived models in lending decisions (Wonderlic 1952). The first consultancy was formed in San Francisco by Bill Fair and Earl Isaac in the early 1950s, and their clients were mainly finance houses, retailers, and mail-order firms.

The arrival of credit cards in the late 1960s made the banks and other credit card issuers realize the usefulness of credit scoring. The number of people applying for credit cards each day made it impossible in both economic and manpower terms to do anything but automate the lending decision. The growth in computing power made this possible. These organizations found credit scoring to be a much better predictor than any judgmental scheme, and default rates dropped by 50% or more—see Myers and Forgy (1963) for an early report on such success, and see Churchill et al. (1977) for one from a decade later. The only opposition came from those like Capon (1982), who argued that "the brute force empiricism of credit scoring offends against the traditions of our society." He believed that there should be more dependence on credit history and it should be possible to explain why certain characteristics are needed in a scoring system and others are not. The event that ensured the complete acceptance of credit scoring was the passage of the Equal Credit Opportunity Acts and its amendments in the U.S. in 1975 and 1976. These outlawed discrimination in the granting of credit unless the discrimination "was empirically derived and statistically valid." It is not often that lawmakers provide long-term employment for anyone but lawyers, but this ensured that credit-scoring analysis was to be a growth profession for the next 25 years. This is still the case, and this growth has spread from the U.S. across the world so that the number of analysts in the U.K. has doubled in the last four years.

In the 1980s, the success of credit scoring in credit cards meant that banks started using scoring for other products, like personal loans, while in the last few years, scoring has been used for home loans and small business loans. In the 1990s, growth in direct marketing led to the use of scorecards to improve the response rate to advertising campaigns. In fact, this was one of the earliest uses in the 1950s, when Sears used scoring to decide to whom to send its catalogues (Lewis 1992). Advances in computing allowed other techniques to be tried to build scorecards. In the 1980s, logistic regression and linear programming, the two main stalwarts of today's card builders, were introduced. More recently, artificial intelligence techniques, like expert systems and neural networks, have been piloted.

At present, the emphasis is on changing the objectives from trying to minimize the chance a customer will default on one particular product to looking at how the firm can maximize the profit it can make from that customer. Moreover, the original idea of estimating the risk of defaulting has been augmented by scorecards that estimate response (How likely is a consumer to respond to a direct mailing of a new product?), usage (How likely is a consumer to use a product?), retention (How likely is a consumer to keep using the product after the introductory offer period is over?), attrition (Will the consumer change to another lender?), debt management (If the consumer starts to become delinquent on the loan, how likely are various approaches to prevent default?), and fraud scoring (How likely is that application to be fraudulent?).

There is a fairly limited literature on credit scoring. Ted Lewis, one of the founders of the industry, wrote a monograph on the practical aspects of credit scoring (Lewis 1992). The proceedings of an international conference on credit scoring appeared in the same year (Thomas et al. 1992), and there are several chapters on credit scoring in a textbook (Hand and Jacka 1998). Mays edited a book with chapters on different aspects of credit and mortgage scoring (Mays 1998), while there are a number of textbooks that look at classification problems in general (Hand 1981, 1997). A series of review articles in journals addressed the techniques used in credit scoring (Rosenberg and Gleit 1994, Hand and Henley 1997, Thomas 1998).

1.4 Philosophical approach to credit scoring

The philosophy underlying credit scoring is pragmatism and empiricism. The aim of credit scoring and behavioral scoring is to predict risk, not to explain it. For most of the last 50 years, the aim has been to predict the risk of a consumer defaulting on a loan. More recently, the approach has been to predict the risk that a consumer will not respond to a mailing for a new product, the risk that a consumer will not use a credit product, or even the risk that a consumer will move an account to another lender. Whatever the use, the vital point is that credit scoring is a predictor of risk, and it is not necessary that the predictive model also explain why some consumers default and others do not. The strength of credit scoring is that its methodology is sound and that the data it uses are empirically derived.

Thus credit-scoring systems are based on the past performance of consumers who are similar to those who will be assessed under the system. This is usually done by taking a sample of past customers who applied for the product as recently as it is possible to have good data on their subsequent performance history. If that is not possible because it is a new product or only a few consumers have used it in the past, then systems can be built on small samples or samples from similar products (see Chapter 12), but the resultant system will not be as good at predicting risk as a system built on the performance of past customers for that product.

There is a parallel development to credit scoring in using scoring approaches to predict the risk of companies going bankrupt (Altman 1968). Although this has provided some interesting results connecting accounting ratios to subsequent bankruptcy, because samples are so much smaller than in consumer credit and because accounting information is open to manipulation by managers, the predictions are less accurate than for the consumer credit case.

The pragmatism and empiricism of credit scoring implies that any characteristic of the consumer or the consumer's environment that aids prediction should be used in the scoring system. Most of the variables used have obvious connections with default risk. Some give an idea of the stability of the consumer—time at address, time at present employment; some address the financial sophistication of the consumer—having a current or checking account, having credit cards, time with current bank; others give a view of the consumer's resources—residential status, employment, spouse's employment; while others look at the possible outgoings—number of children, number of dependents. Yet there is no need to justify the case for any variable. If it helps the prediction, it should be used. There are apocryphal stories of the first letter of the consumer's surname being used—although this is probably a surrogate for racial origin. In the U.K., there has been an angry battle with the data protection registrar over whether information on people who have lived at the same address as the consumer could be used. Such information certainly enhanced the prediction of the scorecard; otherwise, the battle would not have taken place.

This battle makes the point that it is illegal to use some characteristics like race, religion, and gender in credit-scoring systems. (This is discussed in more detail in section 8.8.) A number of studies (Chandler and Ewert 1976) have showed that if gender were allowed to be used, then more women would get credit than is the case. This is because other variables, like low income and part-time employment, are predictors of good repayment behavior in females but poor repayment behavior in the whole population. Yet legislators will not allow use of gender because they believe it will discriminate against women.

Other characteristics, although not legally banned, are not used to predict default risk because they are culturally unacceptable. Thus a poor health record or lots of driving convictions are predictors of an increasing risk to default, but lenders do not use them because they fear the approbation of society. However, checking whether the consumer has taken out insurance protection on credit card debts in the case of loss of employment is used by some lenders and is found to be positively related to chance of default. Thus what are acceptable characteristics is very subjective.

This led to a debate in the early 1980s about the ethics of credit scoring between those (Nevin and Churchill 1979, Saunders 1985) who advocated it and those (Capon 1982) who were critical of its philosophy and implementation compared with the subjective judgmental systems based on credit analysts and underwriters' opinions. The former described the advantages of credit scoring as its ability to maximize the risk-reward trade-off, that it gave managerial control of this trade-off, and that it is efficient at processing applications. They argued that credit scoring reduced the need for credit investigations, but this is not really the case. Seeking a bank reference was on the decline anyway since this allowed the applicant's bank the opportunity to make a counteroffer. Using a credit bureau has become a standard fraud check, so the credit bureau might as well give the information on the consumer's credit history and check the validity of the address. The proponents of credit scoring also advocated its consistency over applicants and stated that it improved the information available on the accounts and the quality of the overall portfolio of accounts.

Those against credit scoring attacked its philosophy and the soundness of its methodology. They criticized the fact that it did not give any explanation of the links between the characteristics it found important and the subsequent credit performance. They argued that there was a complex chain of interacting variables connecting the original characteristics and the performance. It is interesting to note that some of the recent developments in credit scoring, like graphical modeling, outlined in section 12.6, do seek to model such chains. The soundness of the statistical methodology was criticized because of the bias in the sample used, which does not include those who were previously rejected. The appropriate size of a sample was questioned, as was the problem of overriding the systems decisions. Other problems highlighted were the collinearity between the variables and the discontinuities that coarse classifying introduces into continuous variables. Eisenbeis (1977, 1978) also noted these criticisms. The credit-scoring industry has long been aware of these criticisms and either has found ways to overcome any deficiency or allows for them in its decision making. These issues are all addressed in Chapter 8.

The overwhelming point is that in the 20 years since this debate took place, scoring systems have been introduced for many different consumer lending products by many different types of lending organization all over the world, and the results have almost always given significant improvements in the risk-return trade-off. The proof of the pudding is in the eating. That is not to say the debate has gone away completely. For example, the introduction of credit scoring into mortgage lending in the U.K. is even now sparking a debate on the role of the underwriters vis-à-vis the mortgage scorecard in the lending institutions.

1.5 Credit scoring and data mining

Data mining is the exploration and analysis of data in order to discover meaningful patterns and relationships. As with mining for minerals, to be successful you have to know where to look and recognize what is important when you find it. In this case, one is looking for information that helps answer specific business problems. Over the last decade, organizations, especially retailers and banks, have recognized the importance of the information they have about their customers. This is because with the advent of electronic funds transfer at point of sale (EFTPOS) and loyalty cards, they are able to gather information on all customer transactions. Moreover, the growth in computer power makes it feasible to analyze the large numbers of transaction data collected. Telecommunications companies, for example, are able to analyze each night all the calls made by their customers in the previous 24 hours. Moreover, new competitors, substitute products, and easier communication channels like the Internet mean that it is much easier for customers to move to rival firms. Thus it is essential for firms to understand and target their customers. This is why they are willing to spend so much money developing data warehouses to store all this customer information and on data mining techniques for analyzing it.

When one looks at the main techniques of data mining, not surprisingly one finds that they are the ones that have proved so successful in credit scoring. In his review of data mining, Jost (1998) acknowledged this fact. The basic data-mining techniques are data summary, variable reduction, observation clustering, and prediction and explanation. The standard descriptive statistics like frequencies, means, variances, and cross-tabulations are used for data summary. It is also useful to categorize or "bin" the continuous variables into a number of discrete classes. This is exactly the technique of coarse classifying (see section 8.8) that has proved so useful in credit scoring. Trying to find which variables are most important, so one can reduce the number of variables one has to consider, is standard in many statistical applications, and the methods used in credit scoring are also the ones that are successful in the other data-mining application areas. Clustering the customers into groups who can be targeted in different ways or who buy different products is another of the data-mining tools. Credit scoring also clusters the consumers into different groups according to their behavior and builds different scorecards for each group. This idea of segmenting into subpopulations and then having a portfolio of scorecards—one for each subpopulation—is discussed in detail later in the text.

The prediction uses of data mining—who will respond to this direct mailing or who will buy this financial product in the next year—are discussed in Chapter 11, since the techniques used are exactly those developed earlier for credit scoring. Jost (1998) described explanation analysis as a major use of data mining but then went on to make the point that rarely are satisfactory explanations possible. Instead, data mining uses segmentation analysis to indicate which segment is most likely to exhibit a particular type of behavior. The philosophical difficulties that Capon (1982) had with credit scoring still occur in this context in that it is very hard to explain why that segment behaves in that way. Human behavior is complex and it seems all one can do is point out the correlations between the segments and their subsequent purchase and repayment behavior. The idea of causality between the two should not be attempted.

Thus it is clear that data mining has at its heart the techniques and methodologies of credit scoring but these are applied in a wider context. It is be hoped that those who use and advocate data mining study the successes and developments of credit scoring since credit scoring proves a useful guide to the pitfalls to be avoided and the ideas that will work in these other application areas.

This page intentionally left blank

Chapter 2

The Practice of Credit Scoring

2.1 Introduction

In this chapter, we introduce the basic operation and ideas of scoring, at least from a business and lending perspective. Many of these ideas are referred to in later chapters.

First, we set the scene of how credit assessment was carried out before the widespread use of credit scoring. Then we examine at a high level what a scorecard is and how it fits into the lender's overall operation. We touch on the required data and how they are managed. We also look at some of the other parties involved—the credit bureaus and credit-scoring consultancies.

By the end of this chapter, the reader should have a grasp of the basic business framework in which scoring is used so that, as issues and challenges are discussed in later chapters, the context in which these arise can be understood.

2.2 Credit assessment before scoring

Not so long ago—certainly in the 1970s in the U.K. and the U.S. and perhaps, for a few lenders, still in the late 1990s—credit scoring was not used. Traditional credit assessment relied on "gut feel" and an assessment of the prospective borrower's character, ability to repay, and collateral or security. What this meant was that a prospective borrower did not approach a bank manager or building society manager until they had been saving or using other services for several years. Then, with some trepidation, an appointment would be made and, wearing Sunday best, the customer would ask to borrow some money.

The manager would consider the proposition and, despite the length of relationship with the customer, would ponder the likelihood of repayment and assess the stability and honesty of the individual and their character. He—the manager was invariably male—would also assess the proposed use of the money, and then he might ask for an independent reference from a community leader or the applicant's employer. He might arrange for a further appointment with the customer and then perhaps reach a decision and inform the customer. This process was slow and inconsistent. It suppressed the supply of credit because the prospective borrower would have such a relationship with only one lender. If one asked the bank manager about this, he would answer that such a task required many years of training and experience. (Of course, one could not learn from one's mistakes because the process was so cautious that these were rarely made.)

These disadvantages had existed for some time, so what happened to change things? During the 1980s in the U.K., many changes occurred to the lending environment. Some of these changes were as follows:

- Banks changed their market position considerably and began to market their products. This in turn meant that they had to sell products to customers not only whom they hardly knew but whom they had enticed.

- There was phenomenal growth in credit cards. Sales authorizations of this product meant that there had to be a mechanism for making a lending decision very quickly and around the clock. Also, the volumes of applications were such that the bank manager or other trained credit analyst could would not have the time or opportunity to interview all the applicants. Clearly there would be insufficient numbers of experienced bank managers to handle the volume. During the 1980s, a handful of U.K. operations were dealing with several thousand applications each day.

- Banking practice changed emphasis. Previously, banks had focused almost exclusively on large lending and corporate customers. Now consumer lending was an important and growing part of the bank. It would still be a minority part by value but was becoming significant. Banks could not control the quality across a branch network of hundreds or thousands of branches, and mistakes were made. With corporate lending, the aim was usually to avoid any losses. However, banks began to realize that with consumer lending, the aim should not be to avoid any losses but to maximize profits. Keeping losses under control is part of that, but one could maximize profits by taking on a small controlled level of bad debts and so expand the consumer lending book.

There were many other reasons why scoring came to be introduced in the U.K. during the 1980s, although, as we said, some lenders continue to ignore scoring. Most of these are small lenders who cannot or choose not to expand their lending book in a controlled way, and so they do have the luxury of a few well-trained and experienced managers who can make the critical underwriting decisions. For the vast majority of the market, scoring is used in one form or another.

In consumer lending, it is probably used most in credit cards, because of the volumes and 24-hour coverage. Also, once the decision has been taken to grant a credit card, the lending process does not end. Decisions may be required, not only on authorizations but on whether to amend the customer's credit limit or even on whether to reissue a card and for how long.

For installment loans, the scoring decision is a little simpler. The basic decision is whether to grant a loan. In some situations, it may be more complex as there are additional parameters to consider. These might include the interest margin or interest rate to charge, or the security to demand, or the term for which the money is to be borrowed. Other parameters are possible, but the key factor with most loans is that once the lender has agreed to lend the money, provided the borrower makes payments when due, there is no further decision to make.

For overdrafts, the situation is similar to that for credit cards. A decision is required regarding what the overdraft limit should be. If customers operate their accounts within that limit, there is little else to do. However, if a check is presented or an electronic payment request is made that would take the account over the current overdraft limit, then there is a decision on whether to pay the check or to return it unpaid. If the customer requests an increase in the overdraft limit, a decision is required on whether to grant or decline this.

While credit card limits are reviewed from time to time to see if an increase is warranted, this has not yet happened on a widespread basis with overdrafts.

Once we begin to consider some other types of lending, other factors are introduced. Two of the main additional factors are ownership and security. With leasing or hire purchase, the item concerned—typically a car—may not be actually owned by the consumer until the final payment has been made. This may have an effect on the likelihood of the consumer continuing to make payments. On the one hand, it may be easier for the lender to reclaim the outstanding balance by repossessing the car. This is easier than having to go through court action to establish the debt and then have the court order the consumer to repay the debt, often at an even slower schedule than was in the original credit agreement. On the other hand, the consumer may also find it easier once they get into arrears to return the asset. Of course, the matter is not that simple in practical terms. For example, if the consumer has made, say, 27 of 36 payments, there may be no entitlement to a refund when the car is repossessed. For example, the lender may find that when repossessing the car, it is no longer worth the outstanding balance, and so the consumer may still have to be pursued for money. Also, it may be difficult to repossess the car because it cannot be located.

Of course, if repossession does occur, the lender now has to concern itself not only with lending money but also with the value of second-hand cars.

With mortgages, the consumer owns the property. The borrower will take a legal charge over the property. This prevents the property being sold without the debt being repaid. It also gives the lender a right to repossess the property when the arrears situation has deteriorated, provided a prescribed legal process is followed. As with leasing, a key issue here is whether the property will be worth enough on repossession to extinguish the debt. Falling property values can be caused by the external market or by the residents failing to look after the property. On the other hand, in most cases, the mortgage is the last of the consumer's payments to go into arrears as there is a clear wish to protect the home. (This is also referred to in section 13.6.) After all, people need to live somewhere.

2.3 How credit scoring fits into a lender's credit assessment

At a high level, the potential borrower presents a proposition to the lender. The lender considers the proposition and assesses the related risk. Previously, bankers wove some magic which allowed them to gauge whether the risk was of an acceptably low level. With scoring, the lender applies a formula to key elements of the proposition, and the output of this formula is usually a numeric quantification of the risk. Again, the proposition will be accepted if the risk is suitably low.

How credit scoring fits into the lender's assessment may vary a little from product to product. Let us consider, as an example, an application for a personal loan. Nowadays, the applicant will complete an application form. This could either be a paper form or on a computer screen. An application may be made onscreen if the application is made over the Internet or is made within the branch, where data may be entered by either the customer or the branch staff. Screen-based entry may also arise if the lender is a telephone-based lender and the applicant provides details over the telephone, which are then entered into a screen by a telesales operator.

Typically, the application data will be scored. Not all the application data are used in calculating a credit score. However, as we shall see below, other pieces of information are needed for a variety of purposes, including identification, security, and future scorecard development.

The credit score calculation may also include some information from a credit bureau (see section 2.9). In many cases and in many environments, the result of the credit-scoring process makes a recommendation or decision regarding the application. The role of subjective human assessment has been reduced to a small percentage of cases where there is a genuine opportunity for the human lender to add value.

To make matters more concrete, let us consider a simple scoring operation. Suppose that we have a scorecard with four variables (or characteristics): residential status, age, loan purpose, and value of county court judgements (CCJs); see Table 2.1.

Table 2.1. *Simple scorecard.*

Residential Status		Age	
Owner	36	18–25	22
Tenant	10	26–35	25
Living with parents	14	36–43	34
Other specified	20	44–52	39
No response	16	53+	49

Loan Purpose		Value of CCJs	
New car	41	None	32
Second-hand car	33	£1–£299	17
Home improvement	36	£300–£599	9
Holiday	19	£600–£1199	–2
Other	25	£1200+	–17

A 20-year-old, living with his or her parents, who wishes to borrow money for a second-hand car and has never had a CCJ, will score 101 ($14 + 22 + 33 + 32$). On the other hand, a 55-year-old house owner, who has had a £250 CCJ and wishes to borrow money for a daughter's wedding, would score 127 ($36 + 49 + 25 + 17$).

Note that we are not saying that someone aged 53+ scores 27 points more than someone aged 18–25. For the characteristic age, this 27-point differential is true. However, as we can clearly see, there are correlations involved. For example, someone aged 53+ may also be more likely to own their house than someone aged 18–25, while it might be quite rare to find someone aged 53+ living with their parents. Thus what we might find is that someone in the older age category may score, on average, 40 or 50 or 60 points more once the other characteristics have been taken into account.

In setting up a credit-scoring system, a decision will be made on what is the pass mark. This is a simple thing to implement but not necessarily a simple thing to decide on. This is touched on in Chapter 9.

Let us suppose that in the example above, the pass mark is 100. Thus any application scoring 100 or more would carry a recommendation for approval. This would be the case whatever the answer to the four questions. What scoring allows, therefore, is a trade-off, so that a weakness in one factor can be compensated for a strength in other factors.

In assessing an application for credit, the lender collects information on the applicant. This can be from a variety of sources, including the applicant, a credit bureau, and the lender's files of the applicant's other accounts. Credit bureau reports are usually available electronically but paper reports are still available (and are still quite common in business and commercial lending).

The lender will examine the information available and calculate a score. There are many ways to use this score.

Some lenders operate a very strict cutoff policy. If the score is greater than or equal to the cutoff, the application is approved. If it is less than the cutoff, the application is declined.

Some lenders operate a simple variation to this. A referral band or gray area is created. This might be 5 or 10 points on one or both sides of the cutoff. Applications falling into such a gray area are referred for a closer look. This closer look might involve some subjective underwriting or might involve seeking further information that is still assessed objectively.

Some lenders operate policy rules that force potentially accepted cases into a referral band. For example, this might be where the application achieves the cutoff but there is some adverse event in the credit bureau information, e.g., a bankruptcy. In other words, we would not be permitting the strength of the rest of the application to compensate automatically for a weakness.

Some lenders operate what are called super-pass and super-fail categories. We discuss the role of the credit bureau below, but it is acknowledged that getting a credit bureau report incurs a cost. Therefore, certainly there are cases that score so poorly that even the best credit bureau report will not raise the score to the cutoff. These would be classed as super-fails. At the other extreme, we might have cases that score so well that even the worst credit bureau report will not reduce the score to below the cutoff. These are super-passes. In effect, these lenders are operating two or three cutoffs: one to define a super-pass, one to define a super-fail, and, in between, a cutoff to be used once the credit bureau information has been used. An alternative use of this might be where the cost of the credit bureau report is low relative to the cost of capturing the applicant information. This might occur in a telephone-based operation: the credit bureau report might be obtained early on, and if a case scores very poorly the application could be curtailed quickly.

Some lenders operate risk-based pricing or differential pricing. Here, we may no longer have a simple fixed price. Rather, the price is adjusted according to the risk (or the profit potential) the proposition represents. Instead of having one cutoff, the lender may have several. There might be a high cutoff to define the best applicants who might be offered an upgraded product, another cutoff for the standard product at a lower interest rate, a third cutoff for the standard product at the standard price, and a fourth cutoff for a downgraded product. In commercial lending, to some extent, the assessment of price takes the risk into account, although other issues, such as competition and the customer relationship, will also have a bearing.

2.4 What data are needed?

Requirements for data fit several categories. Table 2.2 presents listed some ideas of what data may be required. However, in most credit environments, much of the data collected are needed for more than one purpose.

2.5 The role of credit-scoring consultancies

Once again, things have changed over the past 10 or 15 years. Until the mid 1980s, a lender who wanted a scorecard would approach an external scorecard developer and contract with it. Much of the process of actual scorecard development is covered in later chapters. At a high level, this approach involved the lender providing a sample of its data and the scoring developer producing a model with the data.

There were few scoring developers in each market but they provided a useful service. They tended to have very large computers that had enough space to carry out enormous

Table 2.2. *Reasons for data collection.*

Purpose	Examples
To identify customer	Name, address, date of birth
To be able to contract with customer	Name, address, date of birth, loan amount, repayment schedule
To process/score the application	Scorecard characteristics
To get a credit bureau report	Name, address, date of birth, previous address
To assess marketing effectiveness	Campaign code, date of receipt of application, method of receipt—post, telephone, Internet
To effect interbank transfers of money	bank account number, bank branch details
To develop scorecards	Any information legally usable in a scorecard. (Law may vary from country to country.)

calculations. Also, many of them either had direct links to credit bureaus or were part of a larger organization which itself had a credit bureau.

Things have changed but not completely. There are still credit-scoring companies who develop scoring models. They also advise on strategies and implementation and provide training where required. Some of the advantages of having an external developer remain—the credit bureau's links and also the fact that developing models for many lenders allows the external developer an opportunity to identify trends across an industry or portfolio type.

On the other hand, many sizable lenders have built up their own internal analytic teams who can develop scorecards at a fraction of the cost. The growth in power and the reduction in cost of computers have facilitated this internal development. As well as the benefit of lower costs, internal teams may understand the data better and should be better placed to anticipate some of the implementation challenges. Scoring consultancies have arisen to bridge the gap between internal and external development. These do not do the development work but advise internal development teams and keep them away from many pitfalls, both practical and analytic.

At the same time, the traditional scoring vendors have greatly changed their position. Many of them are now much more willing for the development to be a cooperative venture. They will openly discuss methodology and are willing to receive input from the lender, who is, after all, the customer. This cooperation might produce a better scorecard—better at least in the sense that it will generate fewer implementation challenges. It might also enhance the customer's understanding. And it could generate different types of problem for the consultancy and so advance the knowledge of the industry as a whole.

2.6 Validating the scorecard

In Chapter 8, we discuss a typical scorecard development. However, at the simplest level, when a scorecard is built, it is by necessity built on historical cases. This is because we need to allow these cases some time to mature so that we know their performance. Before using or implementing a scorecard, it is important to validate it. One validation check that is carried out is to compare the profile of the cases used in the scorecard development with the profile of current cases. Differences will occur. These may be due to differences in branding, in marketing, or in action by a competitor, and they must be investigated and understood.

2.7 Relation with information system

Not all data used appear on an application form. Some are obtained from a credit bureau (see section 2.9). However, the bank may extract data from its own files. For example, in assessing a personal loan application, a strong scorecard variable could be the performance of the applicant's previous loan. Even more likely is that for an application for an increased credit card limit or overdraft limit, the recent performance on the credit card account or current (checking) account is highly predictive and will indicate whether the customer's finances are in or out of control.

Also, in developing a bank of data for scorecard monitoring and tracking or future scorecard development, details of the customer's performance on the loan or credit card must be saved. The more detail stored, the greater the opportunity to identify niches and opportunities or to build a better scorecard.

2.8 Application form

In section 2.3 and Table 2.2, we looked at the data we might need to process a lending application. Clearly, the more data we need to get from the applicant, the longer we need to make the application form. The longer the application form, the less likely the applicant is either to submit the application or to complete the details. Therefore, there is often pressure to make the process as simple as possible for the applicant.

One way to do this is to reduce the form to its bare minimum. Unfortunately, this sometimes makes future scorecard development difficult. For example, if we do not capture details on a characteristic such as Time with Bank, we cannot analyze it to discover if it would help to predict future performance.

Another way to do this is to find alternative sources for the same or approximately equivalent information. The credit bureau may be able to supply information rather so the applicant need not be asked. This is particularly the case with lenders that share account information (see section 2.9). It may also be the case that we can use the length of time that someone is registered at an address to vote as a proxy for the length of time they have resided at the address.

2.9 Role of the credit bureau

Credit bureaus (or credit reference agencies) are well established in the U.S. and in the U.K. In other countries in Western Europe and in the major developed countries, they are in varying stages of development. Eastern Europe, for example, has only recently begun to tackle the issue of how to develop them. Where they are well established, they are state owned or there is a small number of very large players in the market. In the U.S., there are currently three major bureaus, while the U.K. has two.

To understand the role of credit bureaus, we can examine how they arose. Prior to their widespread use, when one considered a lending application, one might write to an employer or a bank for a reference. Certainly, in the U.K., these references became more guarded and less useful. In addition, if Bank Ours, for example, became aware that Mr. Brown is applying to Bank Other for a credit card or a mortgage, before replying to the reference, Bank Ours could offer him one of its credit cards or mortgage packages. Of course, the bank reference would reveal only, at best, details of the performance of the bank account, a general indication of the prospective borrower's character, and any adverse information of which the bank or the branch at which Mr. Brown banked was aware. As credit availability was expanded, bad debts began to appear. What was galling was that these bad debts were easily avoidable as indications that they could occur were available at the time of the application.

Before describing how credit bureaus operate in these two nations, we should recognize the very different position that U.S. and U.K. credit bureaus occupy. A principal piece of legislation governing credit bureaus in the U.K. is the Data Protection Act. In the U.S., one of the key pieces of legislation is the Freedom of Information Act. Therefore, the two regimes start from opposite ends of a spectrum. In very rough terms, in the U.S., information is available unless there is a good reason to restrict it. Also, in very rough terms, in the U.K., information is restricted or protected unless there is a good reason to make it available. Thus U.S. credit bureaus have a greater wealth of information on consumers.

In both environments, the credit reference agencies first began to accumulate publicly available information and put it in a central point. Even in the 1980s, this might be achieved with a large room full of index card cabinets. On receipt of an inquiry, an agent would put the inquirer on hold and run up and down the room accessing relevant cards. Obviously, over time, this became computerized. This means that the inquiry can also be made electronically, by a human operator dialing up the inquiry service or, more commonly, two computers talking to each other. This public information might now include electoral roll information and information on public court debts. Using the computing power, the bureaus are also able to link addresses so that, when a consumer moves house, the debt and the consumer do not become detached.

The bureaus also act as agents for the lenders. Lenders contribute to the bureau details of the current status of their borrowers' accounts. These statuses can be viewed and used by other lenders when considering a credit application. They can also be viewed and used by other lenders in marketing although with some increased greater restrictions on the use of the data. (In the U.K., this agency arrangement works for the lenders on a reciprocal basis. Roughly speaking, if the lender contributes only details on their defaulting customers, they get to see only details on other lenders' defaulting customers.)

Another service that bureaus offer is to accumulate details of all inquiries and try to assess mismatches and potential fraudulent applications. Clearly, the greater the level of detail with which to work, the better. Major bureaus also offer a variety of application processing services.

A further service used by many lenders is a generic score. This score is calculated from a scorecard built by the bureau based on its experience with millions of applications and millions of credit history records. It is particularly useful in cases where the lender is not large enough to develop scorecards for their own portfolios or in the early year or two of a new product. It is also used to get an up-to-date view of the borrower's credit position as it will incorporate the borrower's recent credit performance with all contributing lenders and any inquiries being made as a result of fresh credit applications. Indeed, some lenders, especially in credit card portfolios, buy a score for each of their cardholders every month and use these to assess how to deal with cases which miss payments or go overlimit, or when and by how much to increase the customers' credit limit.

Other discussions of the credit bureaus and of the data available from them appears in sections 8.5 and 13.3.

2.10 Overrides and manual intervention

For a variety of reasons, the decision or recommendation offered as a result of credit scoring an application for credit is not always the one followed. First, there is the fact that the applicant may appeal. In the U.K., the *Guide to Credit Scoring* (Finance and Leasing Association 2000) encourages lenders to have a process that allows this. Lenders need to consider these appeals, especially as there may have been a wrong decision, e.g., due to the incorrect keying of application details or incorrect information loaded at the credit bureau.

Second, we talked earlier about the fact that some lenders will operate a gray area or referral band, above or below or surrounding the optimal cutoff. Therefore, case which reach a final decision to accept or decline having gone through an intermediate step of referral might also be appropriate for consideration as overrides.

A possible third reason may arise when we develop a credit scorecard for applications for a product, but then, in making a decision on whether to lend, we take a customer rather than account view. For example, we may consider a loan application as a marginal decline but recognize that the applicant also has a large amount of money tied up in a savings account as well as family and business connections. Further, if these other connections generate profit and we consider that these connections might be in jeopardy were the loan application declined, we may overturn the decline decision. Now such claims are quite common in large lending organizations. However, it needs to be recognized that few institutions have the data and systems to be able to accurately assess the profitability or otherwise of a connection, especially when we must project that connection's worth to the lender into the future. The subject of assessing customer and connection profitability is too large a topic for this book. Indeed, it could occupy a text on its own. However, the reader must recognize it as an issue to be addressed.

For whatever reason, overrides are created. They can be of two types, i.e., in two directions. First, we may have cases that are acceptable to the scorecard but that are now declined. These usually will be of marginal profitability and so they will not hugely affect the profitability of the portfolio. However, they may affect our ability to assess the accuracy of the scorecard.

We may also have cases that were not acceptable on the scorecard but that have been approved. These are more serious because they may be accounts on which we expect to lose money and we are adding them to the portfolio. This topic is further explored in section 8.10.

2.11 Monitoring and tracking

Monitoring and tracking are covered in later chapters, especially Chapter 9. However, it is worthwhile to introduce the item now and to establish some key concepts.

Many practitioners use these two terms—tracking and monitoring—almost identically and interchangeably. In this text, we differentiate between these two activities. Monitoring is passive, rather like traffic census takers who sit beside a road and mark off on sheets the number of different types of vehicle that pass a particular point in discrete intervals of time. On the other hand, tracking is active, following one's quarry until either they are captured or the tracker discovers their base.

In scoring, monitoring a scorecard is a set of activities involved in examining the current batch of applications and new accounts and assessing how close they are to some benchmark. Often this benchmark is the development sample, although it need not be. In scoring, tracking involves following groups of accounts to see how they perform and whether the scorecard's predictions come to pass.

Most monitoring activity for a tranche of business can be carried out soon after the lending decisions have been made. It may include a forecast itself, but this forecast, of provisions, attrition, write-off, etc., is made soon after the point of application.

Tracking activity may take up to two years after this point. At a variety of points in the life of a tranche of accounts, we should assess whether this particular group of accounts is performing as the scorecard suggests that it should. If we can identify early on that they are performing worse, we may be able to take corrective action and mitigate any losses. On the other hand, if the business is performing better, there is the opportunity to expand credit assessment and to take on additional profitable business.

Earlier, we introduced the idea of credit-scoring consultancies. A service that they can provide is to do the lender's monitoring and tracking. This might be attractive for a small organization where senior management wish to benefit from a scoring approach but without the overhead of an analytic team.

2.12 Relationship with a portfolio of lender's products

As has been touched on, in reaching a lending decision, we may include details of a customer's other holdings. These may be because they are credit related. An example is the case in which the customer has an existing credit card and has applied for a personal loan. We may also include details because they are important for the management of the customer relationship and its profitability. Examples would be cases in which the customer has applied for a personal loan and already has a mortgage or a substantial deposit.

There is also the issue of defining the customer and the boundaries of the customer relationship. We may be trying to assess a small business loan application from a small accountancy partnership that repeatedly refers good potential new customers. Similarly, we may be about to decline a loan application from someone whose extended family represents highly profitable business to the lender but may consider moving their connections or their business. When stepping beyond a simple consideration of one product for one consumer—as we should in managing our business—many issues require careful consideration if we are to strike out in the general direction, at least, of a position of optimum profitability.

Chapter 3

Economic Cycles and Lending and Debt Patterns

3.1 Introduction

In this chapter, we explain how changes in macroeconomic factors affect the demand and supply of credit and how changes in the demand and supply of credit combine with other factors to affect an economy's output. This will allow us to understand the profound role that credit has in the workings of a nation's economy. From a strategic point of view, we are then able to understand how and when the demand for credit is likely to change. We also explain how the output of an economy affects default rates over time. In this chapter, we first describe the cyclical variations in credit over time. Second, we explain the relationships between the volume of credit and variations in output of an economy. Third, we explain how the state of the economy affects default behavior.

3.2 Changes in credit over time

Figure 3.1 shows quarterly data for the U.K. for net consumer credit extended (new credit extended less payments) and gross domestic product (GDP), a measure of the output of an economy and the total income of its participants, for the period 1965Q1–1999Q2. Figure 3.2 shows figures for consumer credit outstanding between 1976Q1 and 1999Q2. All series are measured at constant (1990) prices. Between 1963 and 1988, the growth rate of net credit extended per quarter was much greater than that of GDP, and the same is true for debt outstanding between 1976 and 1990. However, between 1988Q3 and 1992Q2, net credit extended per quarter plummeted and then rose again between 1992Q3 and 1994Q3. The corresponding changes in the stock of debt outstanding can be seen with a decrease in its growth rate from mid 1988 until 1992.

However, these figures do not clearly show cyclical variations in the volumes. To achieve this, we need to remove the trend. This has been done, and the results are shown in Figure 3.3, which is reproduced from Crook (1998). Crook argued that a number of conclusions are suggested. First, by identifying the dates of peaks and troughs it seems that on average those in net credit extended lead those in GDP by +2.9 quarters. That is, people take more credit (net of payments) approximately nine months before the domestic economy makes the items or provides the services they wish to buy and, conversely, people take less credit (net of payments) approximately nine months before the home economy reduces production of the goods and services they wish to buy. Second, Crook found that

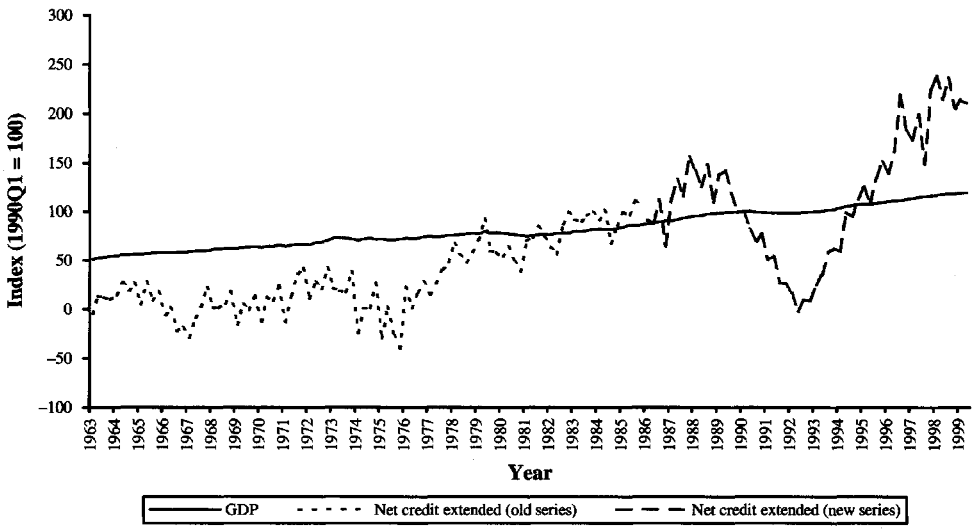


Figure 3.1. *U.K. net consumer credit extended.*

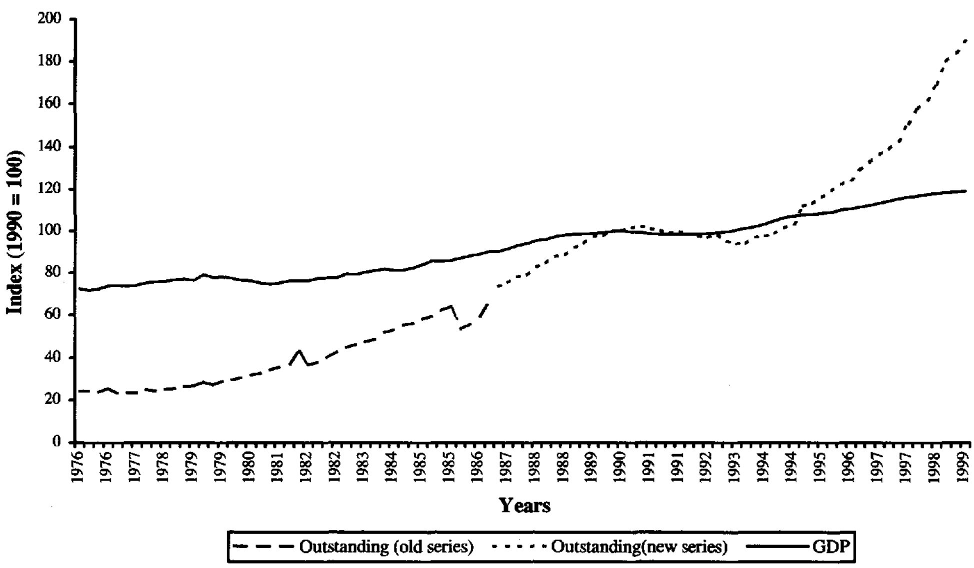


Figure 3.2. *U.K. consumer credit outstanding.*

the average length of the GDP and net credit extended cycles was almost identical at 18.67 and 18.25 quarters, respectively.

Figure 3.4 shows similar figures for the U.S. The data have been manipulated to reveal turning points rather than magnitudes. Figure 3.4 suggests that, using data that when detrended cover the period 1949Q4–1996Q2, the turning points in consumer credit debt outstanding follow those in GDP. The average number of quarters by which outstandings lead GDP is 1.8, and the range of leads is from -1 quarter to $+7$ quarters. Figure 3.4 suggests that the turning points in net credit extended lead those in GDP by typically one or two quarters, the average being $+1.25$ quarters.

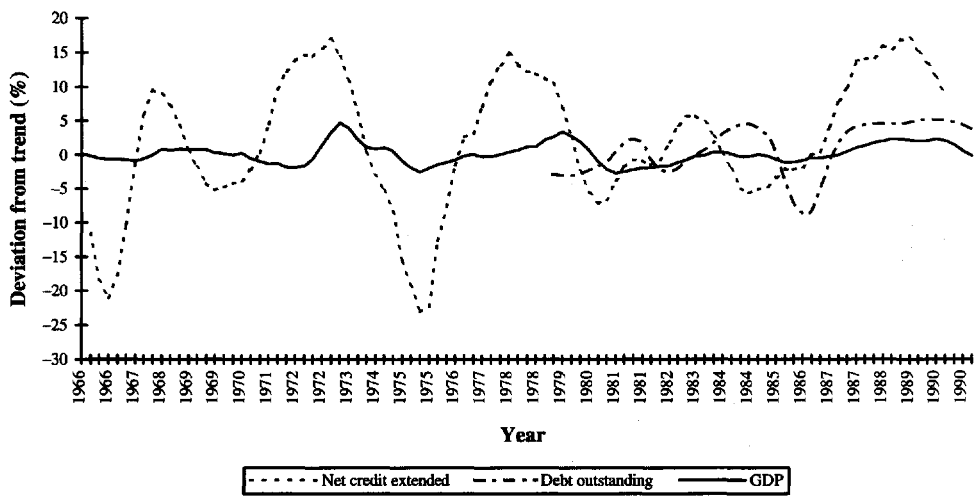


Figure 3.3. *U.K. cyclical behavior of GDP, net consumer credit extended, and consumer debt outstanding. (Reproduced from Crook* (1998) *with permission.)*

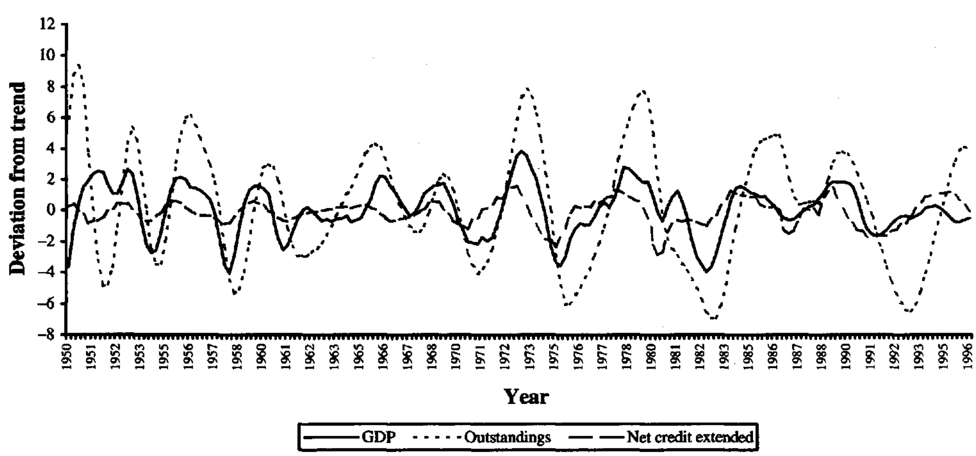


Figure 3.4. *U.S. cyclical behavior of GDP, net consumer credit extended, and consumer debt outstanding.*

Thus far, we have considered some apparent correlations in credit data. To understand how the volume of credit and output relate to each other over time, we need to consider some economic explanations of the relationship and we need to know the results of far more detailed correlation analyses that have tested these explanations. This is our task in the remainder of this chapter.

3.3 Microeconomic issues

3.3.1 Present value

We use the concept of present value often in this book—the first time in this section. Consider an individual investing, today (period 0), $\$R_0$ in a savings account that gives an interest rate of r per year. In year 1 the individual would receive $R_1 = R_0(1 + r)$. When we discount we ask the question, How much does one need to invest today to receive R_1 in year 1, given the interest rate of r? We know the answer is R_0, which can be derived simply as $\frac{R_1}{(1+r)}$. Now the individual may gain receipts in each of several years, so to calculate the amount the person would need to invest today to receive all these amounts, we simply multiply each future receipt by $\frac{1}{(1+r)^t}$, where t is the number of years into the future in which the receipt is gained. (The term $\frac{1}{(1+r)^t}$ is known as the discount factor.) Therefore, the amount a person would need to invest today (year 0) to gain R_t, for each of $t = 1, 2, \ldots T$, which is called the present value of these future receipts, is

$$PV = \frac{R_1}{(1+r)} + \frac{R_2}{(1+r)^2} + \cdots + \frac{R_T}{(1+r)^T}. \tag{3.1}$$

If, in order to gain these future receipts, we need to spend C_0 in the current period, the present value of these future receipts, net of the current outlay, called the net present value (NPV), is

$$\text{NPV} = -C_0 + \frac{R_1}{(1+r)} + \frac{R_2}{(1+r)^2} + \cdots + \frac{R_T}{(1+r)^T}. \tag{3.2}$$

The discount rate, r, is the rate of return on the best alternative use of funds. If instead of investing funds we borrow them, then the discount rate is the lowest borrowing rate instead of the highest lending rate. Notice that discounting has nothing to do with inflation. When we discount we simply remove the interest we would have earned had the future receipt been available for investment today.

3.3.2 Economic analysis of the demand for credit

To understand the effect of the level of output (and thus income) of an economy on the demand for credit we need to consider how the income of an individual borrower would affect his demand for credit. We consider an individual who, for simplicity, is considering how much to consume in each of two periods: periods 0 and 1. We assume the individual wishes to maximize his satisfaction over the two periods but is constrained by his income. First we consider the individual's income constraint.

Suppose the individual receives a certain income in each time period, Y_0 in period 0 and Y_1 in period 1. Suppose the individual can save an amount S_0 in period 0. If we denote his consumption in period 0 as C_0, we can write

$$S_0 = Y_0 - C_0. \tag{3.3}$$

In period 1 the amount of income a consumer has available to consume is his income in period 1, Y_1, plus his savings from period 0, including any interest on them:

$$C_1 = S_0(1+r) + Y_1, \tag{3.4}$$

where r is the rate of interest. Substituting (3.3) into (3.4), we can see that $C_1 = (Y_0 - C_0)(1+r) + Y_1$, which, after manipulation, gives

$$C_0 + \frac{C_1}{(1+r)} = Y_0 + \frac{Y_1}{(1+r)}. \tag{3.5}$$

In other words, the present value of consumption over the two periods equals that of income. Equation (3.5) is known as the budget constraint and can be represented as line AB in Figure 3.5.

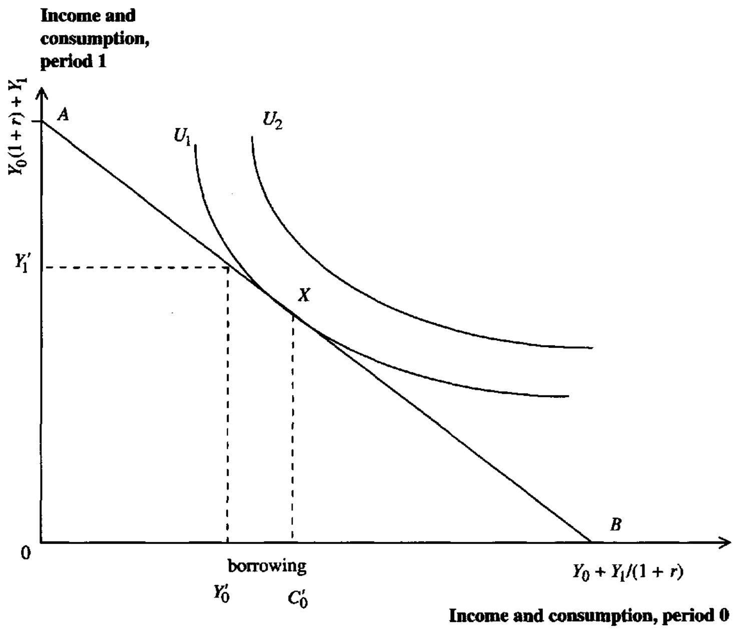


Figure 3.5. *Utility maximization.*

If the individual saved all of his income in period 0, he would receive income of $Y_0(1+r)+Y_1$ in period 1 and would be able to consume this amount then. This is point A in Figure 3.5. Alternatively, the individual might choose to receive income of $Y_0 + \frac{Y_1}{(1+r)}$ in period 0, where $\frac{Y_1}{(1+r)}$ is the amount the individual could borrow in period 0 such that his payment in period 1 just equalled his income, then Y_1. For each dollar by which our consumer increases his consumption in period 0, he must reduce consumption in period 1 by the payment: $(1+r)$. Thus the slope of AB is $-(1+r)$.

Now consider the wishes of the consumer. Suppose he requires additional consumption in period 0 to compensate for loss of consumption in period 1. Then a line representing combinations of consumption in the two periods, each giving equal satisfaction, would slope downward. If we also assume that as the person gains more consumption in period 0 he is willing to give up less consumption in period 1 to compensate, then the line is convex downward, as in the figure. Such a line is called an indifference curve and we have drawn several in the figure. Curves further from the origin represent greater satisfaction because they represent more consumption in period 1 given a level of consumption in period 0.

The combination of consumption in both periods that maximizes the consumer's satisfaction is shown by point X because this is on the highest indifference curve possible, given that the point must also be on the budget line constraint. Given our consumer's incomes of Y_0^1 and Y_1^1, point X indicates that he wishes to consume more than Y_0^1 in period 0 and less than Y_1^1 in period 1. In short, he wishes to borrow an amount $Y_0^1C_0^1$ in period 0.

Now we consider the effect of an increase in income. If the income of the consumer increases, a budget line shifts out parallel to itself because its slope depends only on the interest rate, which has not changed. This is shown in Figure 3.6.

Suppose the additional income is received only in period 0 and the new budget line is $A'B'$ with the new equilibrium at Z. Whether the new amount of desired borrowing, B_2, is larger or smaller than the original amount clearly depends on the shape and position of the

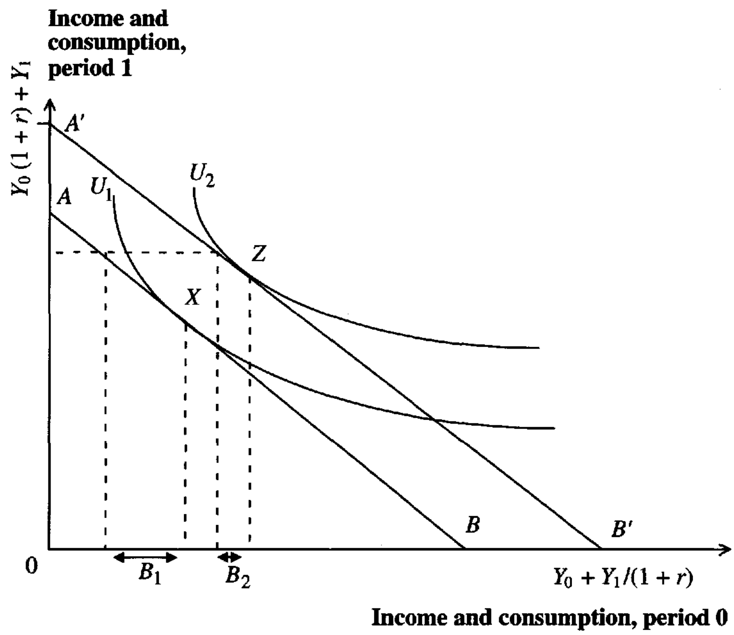


Figure 3.6. *A change in income.*

indifference curves. It may be higher or lower—we cannot say. If the income is received in period 1 this conclusion is reinforced. We conclude that if over time the income of a household rises, the demand for credit may rise or fall.

Equation (3.5) showed that the present value of consumption in the two periods equals an individual's present value of income. It can be extended to as many periods as a person expects to live. Modigliani (1986) argued that given that this is true and that people can borrow or lend as much as they wish, given their lifetime profile of income, each person will typically borrow early in their adult life (when expenditure exceeds income), then save in middle age, and finally, in retirement, run down their savings. This is known as the life cycle theory of consumption. According to this theory we would expect that an increase in the proportion of the population in younger age groups (but above the legal age necessary for acquiring debt) would be associated with increases in the stock of debt. Park (1993) suggests that over the period 1960–1990 the percentage of the U.S. population in the 20–34 years age group increased, especially in 1960–1980. He suggests this may explain the rise in debt outstanding especially in the 1970s and the stabilization of the volume of debt in the 1980s and early 1990s when the percentage of the population in these age groups reached a plateau. Of course, such demographic changes cannot explain short-term changes in demand.

Now consider an increase in the interest rate. This case is shown in Figure 3.7. The slope of the budget line becomes steeper (remember that the slope is $-(1+r)$) and changes from AB to CD. Notice that the new and the old budget line both go through point E, the combination of incomes in periods 0 and 1. This combination is still available after the change.

The initial consumption was at W and the final point is at X. The consumer's desired stock of debt changes from B_1 to B_2, an increase. Economists can show that this is the normal case: when the interest rate rises, the consumer will choose a lower stock of debt.

It has also been argued (see Park 1993) that many individuals have both financial

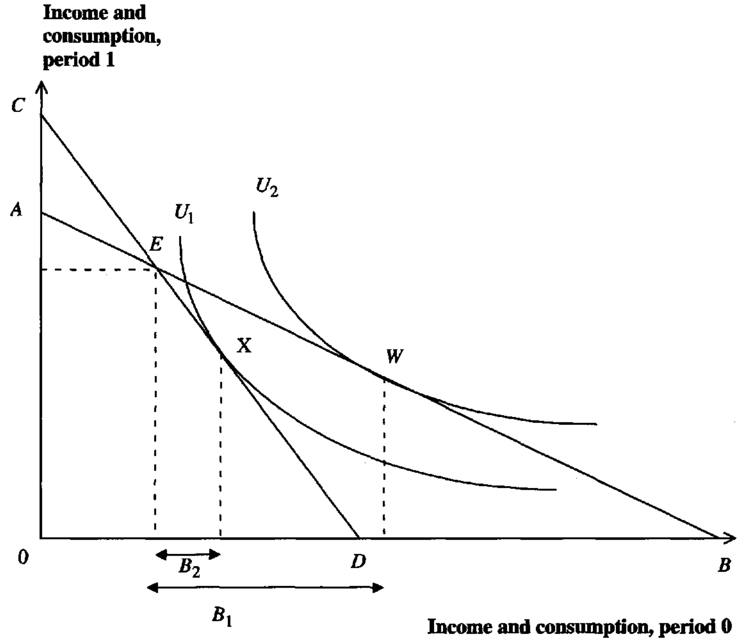


Figure 3.7. *A change in interest rate.*

assets and debt. Although the interest rate on debt exceeds that on assets, people prefer to borrow to pay for some purchases because of the flexibility that having financial assets provides. However, if the difference between the interest rate on assets and that on liabilities changes, so will the proportion of people who take credit rather than run down their assets to buy goods. By examining visually the time series of the difference between these two interest rates and the ratio of net credit extended to consumption expenditure over the period 1972–1992, Park found support for this view.

3.3.3 Credit constraints

If the consumer credit market was very competitive, one might think of representing demand and supply as shown in Figure 3.8.

Given all the factors that affect supply, one might think that the higher the interest rate the more net credit banks would be willing to supply (it is more profitable) and, as we explained above, the less net credit consumers would demand. We would actually observe the equilibrium amount of net credit extended of C_e at an interest rate determined by supply and demand, of r_e. If the interest rate is above r_e suppliers of credit are willing to supply more than consumers wish to borrow, and thus they would offer credit at lower interest rates so that the amount they wish to supply equals the amount consumers wish to borrow. The reverse would happen if the rate was below r_e. Changes in income over the business cycle would shift the demand curve and so affect the equilibrium interest rate and the observed amount of net credit extended. Unfortunately, the market for credit is not so simple.

First, there is evidence that many consumers do not receive all the credit they desire (see, for example, Crook 1996, Jappelli 1990, Crook 2000). Evidence suggests that approximately 18% of households in the U.S. have applied for credit in any five-year period and have been

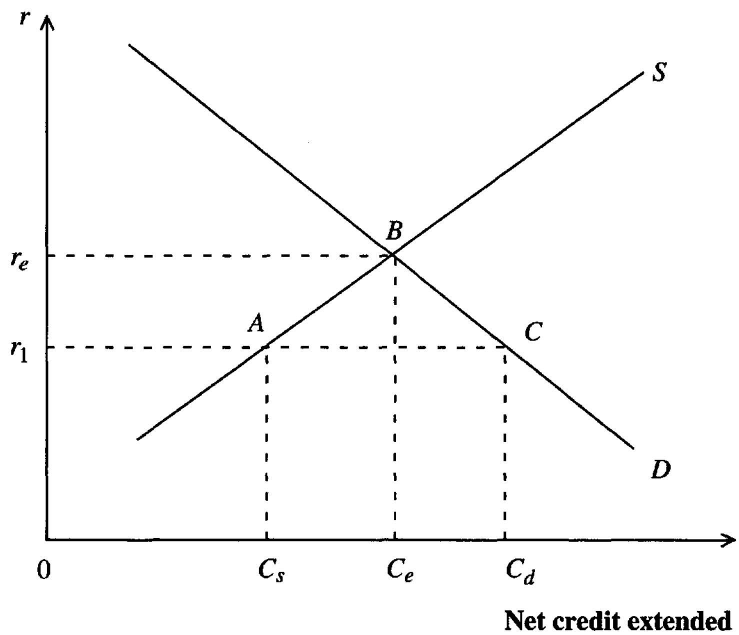


Figure 3.8. *Demand and supply of credit.*

unable to receive all they wished. This proportion varies slightly over time but on the whole is quite stable.

This might be incorporated into our diagram by suggesting that the interest rate is not determined by the intersection of demand and supply but is instead determined by other factors, perhaps the central bank, and the amount of net credit extended that we actually observe is the amount banks wish to offer at this rate. This amount and rate would be r_1 and c_s in the figure.

However, Stiglitz and Weiss (1981) considered four definitions of credit rationing. The first is the type above. Second, some applicants and banks differ in their assessment of the appropriate interest rate to be received by the bank to compensate it for its perception of the riskiness of lending to the applicant. Third, there may be no difference in the rates described in the second definition, but the agreed risk is so high that there is no interest rate that would compensate the lender and at which the applicant would wish to borrow any amount. Fourth, some people may receive credit when others who are identical do not. This may arise as follows. Suppose a bank chose a higher interest rate. This would increase the revenue from each volume of loans. But it also makes each loan more risky because it increases the chance of default. Increasing the interest rate further would result in low-risk applicants turning to other suppliers or not borrowing at all, and so only high-risk borrowers would remain and they yield lower profits. Given that the level of profits and the amount the bank wishes to supply are positively related, a backward bending supply curve is deduced, as shown in Figure 3.9.

Credit rationing can now be seen. If the demand curve is D_0, the equilibrium interest rate is r_0. If the demand curve is D_1 the optimal rate for the bank is r_1, which maximizes its profits, with supply C_1. But the volume of net credit demanded is C_2 at this rate and C_1C_2 is unsupplied. If the business cycle leads to fluctuations in the position of the demand curve, the amount of rationing will also vary over time.

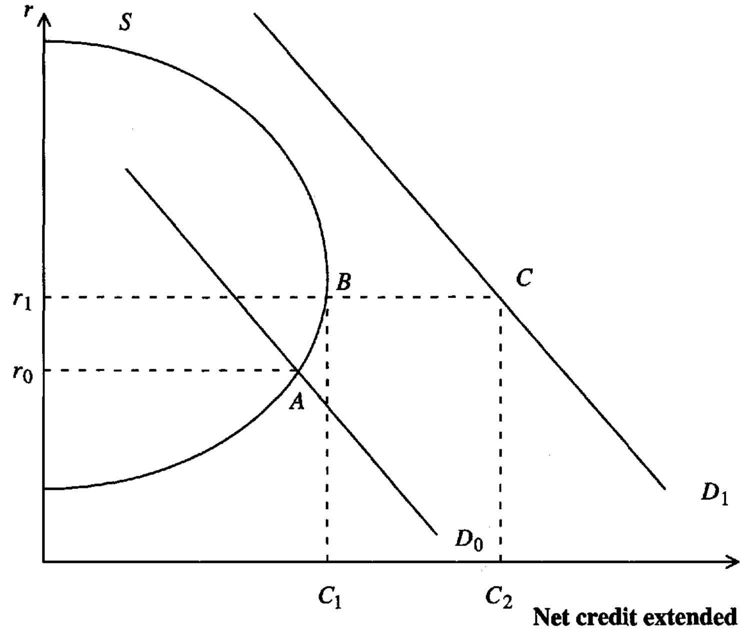


Figure 3.9. *Credit rationing.*

Stiglitz and Weiss proposed their argument as applying to borrowing by firms. However, Drake and Holmes (1995, 1997) found evidence for its applicability to the U.K. mortgage market and in the U.K. market for consumer bank loans, while Martin and Smyth (1991) did likewise for the U.S. mortgage market.

3.3.4 Empirical evidence

We divide the empirical evidence into, first, that relating to the demand for nonmortgage loans and, second, that relating to the demand for mortgage finance.

Table 3.1, relating to nonmortgage loans, shows the results of econometric studies that have related the demand for either net credit extended or new credit extended to various explanatory variables. The elasticity of demand is the percentage change in credit extended when the explanatory variable changes by 1%. Mathematically, the elasticity of y with respect to x is $\frac{\partial y}{\partial x}.\frac{x}{y}$, where $y = f(x)$. Short-run elasticities, that is, elasticities that are estimated from an equation relating quarterly changes in demand to quarterly changes in the explanatory variables, are positive and, generally speaking, have a value approximately equal to one. Thus a 1% increase (decrease) in real personal disposable income (PDI) results in a 1% increase (decrease) in credit demanded. Long-run elasticity, that is, the elasticity estimated from an equation that represents the stable relationship between demand and its explanatory variables, is a little larger. All these studies take the existence of credit rationing into account to identify the demand function. Of particular interest to us is the lagging structures of the affect of income on demand. One of the most sophisticated of the studies is that by Drake and Holmes (1995) using data for the U.K. They established that current values of net credit extended are correlated with current values of income (and other variables) and from this deduced the long-run elasticities. Second, they estimate the relationship between quarterly

Table 3.1. *Elasticities of demand for nonmortgage consumer loans.* (Note: *All four studies relate to the U.K.*)

	Drake and Holmes (1975)	Garganas (1995)	Crook (1989)	Hartropp (1992)
	Net credit extended	New credit extended	New credit extended	Net credit extended
	1977–1992	1957–1971	1977–1985	1968–1987
PDI	+1.349	+0.48 to +1.41	+1.73 to +1.02	positive
Real interest rate				−0.75
Nominal interest rate		−0.20 to −0.22		−0.58
Real net financial wealth	+1.099			
Vacancies		+0.26 to 0.37	+0.07 to +0.04	
Real housing wealth	+1.621			

changes in net credit extended about this long-run relationship and lagged net changes in income. After deleting insignificant variables they find that the change in demand is related to the current change in PDI, the change in housing wealth lagged two quarters, and the change in financial wealth lagged four quarters. They also found evidence of a backward-bending supply curve.

Table 3.2 shows estimated elasticities of demand for mortgage loans. Unfortunately, no clear conclusions concerning the sizes of elasticities can be drawn because the estimates differ markedly between the two studies cited. Possible reasons for the difference include differences in the econometric methodology, the different markets considered with associated differences in tax regimes and other regulations (U.K. versus U.S.), and different methods for taking account of rationing.

Table 3.2. *Elasticities of demand for mortgage debt.*

	Drake and Holmes (1997)	Martin and Smyth (1991)
	(U.K.)	(U.S.)
	1980–1992	1968–1989
PDI	+1.555	+4.860
Nominal interest rate	−0.007	−4.254
House price inflation	+0.031	
Nonprice rationing	+0.056	

3.4 Macroeconomic issues

In the previous section we considered the market for credit alone. Such an analysis is know as a partial equilibrium analysis. In this section we consider the relationship between credit and aggregate output of an economy taking into account the interrelationships between different parts of the economy. When we consider many markets together we are undertaking a general equilibrium analysis.

Most economists would agree that changes in the money supply result in changes in the total output of an economy in the short run. There is a considerable amount of evidence to support this, beginning with Friedman and Schwartz (1963). However, there is considerable debate about *how* a change in the money supply effects changes in real variables (such as output and employment) and the role of credit in this relationship. The mechanism by which changes in the money supply effect changes in real variables is known as the transmission mechanism. While there is no universally accepted mechanism, there appear to be three major views: the money channel, the exchange rate channel, and the credit channel. The role of credit differs between the first two and the third. In this section, we first explain a very simplified Keynesian-type model of the economy. Then we explain each channel, and finally we evaluate the empirical evidence to see if it supports the existence on one channel rather than the others.

3.4.1 A simplified Keynesian-type model of the economy

For simplicity we assume that the price level in an economy is fixed and that the economy is closed: it has no trade with other countries. While these assumptions are unrealistic, they help us to explain the mechanism most simply.

Such an economy consists of many markets, but we assume that they can be represented by two: the market for goods and services and the market for money. In the goods market the total demand for goods and services consists of the planned (or desired) consumption expenditure by households, the planned investment expenditure by firms, and planned government expenditure. We assume that firms always have spare capacity so that the total output of an economy is determined by the volume of goods and services demanded.

What determines planned investment and planned consumption expenditures? We assume that planned investment by firms depends (negatively) on the interest rate they would forego if they used their funds in this way, or the interest rate they would have to pay if they had to borrow funds. This can be justified by assuming that firms invest in those items—plant, equipment, advertising campaigns, etc.—that give a positive net present value (NPV). If the interest rate they forego rises, the NPV of the planned expenditure on plant decreases, and if the rate rises sufficiently the NPV will become negative and the firm would not make the investment. Second, we assume that planned consumption expenditure depends positively on income but negatively on the interest rate consumers have to pay if they borrow funds to spend, for example, on consumer durables or houses. Again, the higher the rate the lower the present value of future satisfaction in each year. Also, the more expensive credit is, the less credit people will wish to take, as we explained in section 3.3. Third, we assume that planned government expenditure is determined by political considerations and is unrelated to output or interest rate. We can represent these relationships by the following equations:

$$I_p = a - br, \tag{3.6}$$

$$C_p = c - dr + eY, \tag{3.7}$$

$$G_p = \overline{G}, \tag{3.8}$$

$$\mathrm{AD} = I_p + C_p + \overline{G}_p, \tag{3.9}$$

where I_p denotes planned investment, C_p denotes planned consumption expenditure, G_p denotes planned government expenditure, Y denotes income and value of output, r denotes the interest rate, AD denotes aggregate demand, and a, b, c, d, e, and $\overline{G}$ all denote constants. In equilibrium, aggregate demand equals actual output:

$$\text{AD} = Y. \tag{3.10}$$

Therefore, we can deduce that

$$Y = \frac{\left(a + c + \overline{G}\right) - (b + d)r}{(1 - e)}, \tag{3.11}$$

which can be written as

$$Y = \theta - \phi r, \tag{3.12}$$

where

$$\theta = \frac{\left(a + c + \overline{G}\right)}{(1 - e)}$$

and

$$\phi = \frac{(b + d)}{(1 - e)},$$

which is known as the IS (investment and saving) curve.

We now consider the market for money. There are several measures of the supply of money. M0 consists of notes and coins in the banking system as reserves held by banks, plus cash outside the banks, such as that held by individuals and firms. M1 consists of cash outside the banks plus sight deposits. Sight deposits are deposits against someone's name, which they can withdraw whenever they wish without any warning, for example, deposits in checking accounts. At this stage we suppose that the supply of money is determined by the central bank, such as the Federal Reserve.

Returning to the demand for money, we suppose that there are only two types of asset in this economy: money and bonds. The term *bonds* is generic and refers to all assets that may give a rate of return: government securities, houses, debt that you are owed, and so on. Given the wealth that each person has, he will keep some of it as bonds and some as money. People, firms, and organizations demand a certain amount of money for several reasons. First, salaries are not perfectly synchronized with expenditures, so each of us wishes to keep some of our wealth as money rather than bonds to enable us to buy things between the times we are paid. The more things we buy, the more money we need. Hence economists assume that the demand for money for this, the transactions motive for holding it, is positively related to output in the economy, Y. But since keeping cash and noninterest bearing bank accounts (checking accounts) gives no (or minimal) interest, whereas bonds do give interest, the higher the interest rate on bonds, the less money we wish to hold to facilitate transactions. A second reason people hold money is to enable them to make purchases they were not planning to make: to repair the car that broke down, to buy an item on sale. Again the amount of money demanded for this precautionary motive is assumed to be related to output. Third, people hold money as an asset. For example, people may put money into a low-risk bank savings account because holding a bond involves a risk that the return will be less than expected. Holding a mix of bonds and money enables each of us to reduce the risk we incur compared with that incurred if we held only bonds. The greater the return on bonds, relative to that in money accounts, the less money people will wish to hold to reduce their risk. In summary we can say that the demand for money is positively related to output and negatively related to the interest rate.

The equilibrium level of the interest rate is determined by the demand and supply of money. Given the level of output, Y, Figure 3.10 plots both the amount of money demanded and the amount supplied at each interest rate. The amount demanded equals the amount supplied at interest rate r_e. To see how the interest rate would adjust if it were not r_e, suppose the interest rate is r_0. At this rate, the supply of money is greater than the amount demanded. Thus, since the alternative to holding money is to hold bonds, at r_0 the demand for bonds must exceed its supply. The price of bonds will rise and the interest rate on them fall.[1] This will continue until there is no more excess demand for bonds and so no more excess supply of money. Thus the interest rate in Figure 3.10 will fall until r_e.

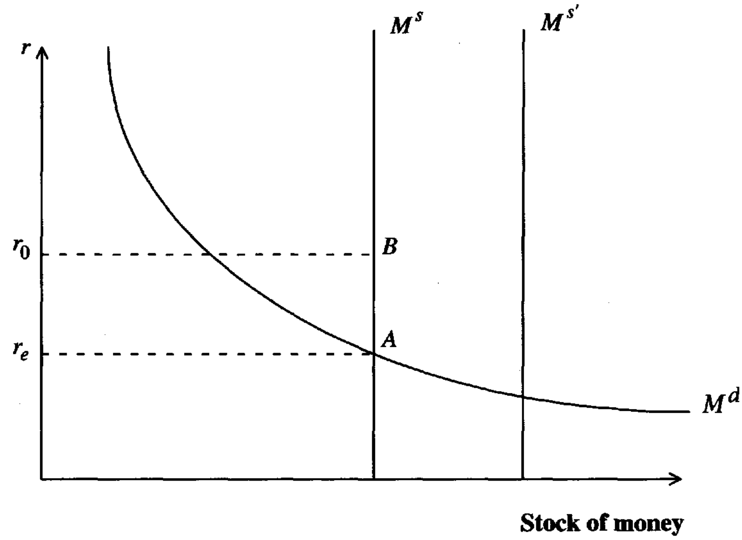


Figure 3.10. *Determination of the interest rate.*

Now suppose that the number of transactions in the economy increased because the level of output increased. The demand for money at each interest rate would increase to facilitate the additional transactions. Thus the demand for money curve would shift to the right and the equilibrium level of interest rate would increase. Thus we can represent the relationship between the level of output, Y, and the interest rate, r, given the money supply as

$$Y = \delta + \lambda r. \tag{3.13}$$

This is known as the LM curve (where LM stands for liquidity preference—money).

We now put the two markets in our economy together, as in Figure 3.11. For the economy to be in equilibrium there must be equilibrium in both the goods and services market and in the money market. The same interest rate and output must exist in both markets. Therefore, the economy is in equilibrium when both curves cross; that is, they are simultaneous equations. The equilibrium values in the figure are r_e and Y_e. In the next sections we explain how changes in monetary variables are related to changes in real variables like output in the economy: the transmission mechanism. The role of credit in a macroeconomy is explained by its role in this mechanism.

[1]For a bond with given payments of R_t in year t, the amount the marginal buyer is just willing to pay for it, its price P, is the present value of these receipts over the life (T years) of the bond:

$$P = \sum_{t=0}^{T} \frac{R_t}{(1+r)^t},$$

where r is the interest rate. Thus r and P are inversely related.

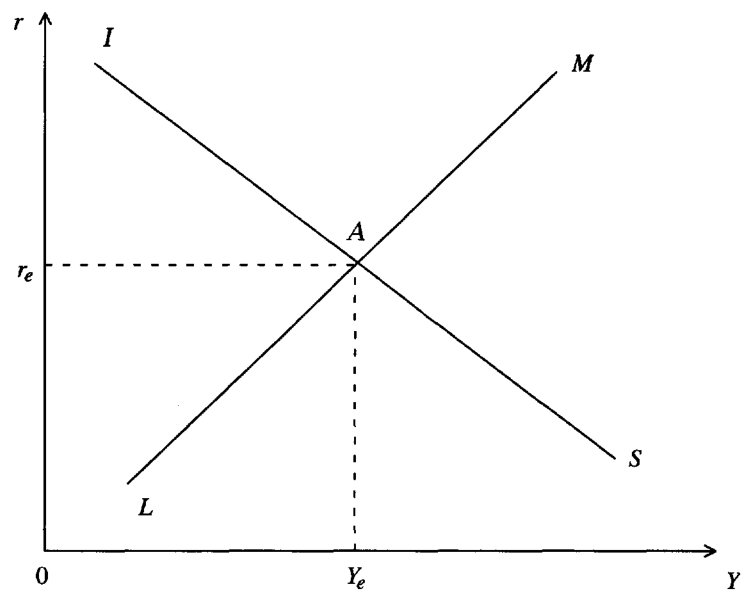


Figure 3.11. *Equilibrium in the goods and money markets.*

3.4.2 The money channel

The money channel of the transmission mechanism, sometimes called the interest rate channel, is the standard Keynesian mechanism explained in most macroeconomics textbooks. Suppose the central bank, the Fed, buys bonds and so pays cash to the private sector—citizens and firms. The private sector will deposit this cash into commercial banks. Each bank's balance sheet consists of assets and liabilities, each of which must sum to the same value. The banks' assets have increased because the banks have received the cash, and their liabilities have increased because the banks have created deposits against the names of those who paid the cash in. If banks choose a cash:deposits ratio of, say, α (such as 3%), they will create deposits of $\frac{\text{cash}}{\alpha}$. This will be done by lending to people who wish to borrow. The borrowers will spend their loan and the recipients of these expenditures will pay their receipts back into banks. In turn, the banks will create deposits against the depositors' names. These receipts will be lent again, and deposited again, and so on. The process will stop when the chosen cash:deposit ratio is reached, but with more cash now in the banks (the original amount plus the additional deposit that came from the Fed). Thus on the assets side of banks' balance sheets is the original cash paid in, plus the value of a number of loans. On the liabilities side are total deposits created, as money is paid in at each round of the process.

Remember that M1 is cash outside the banks plus the sight deposits created. After the cash was paid in it was repeatedly lent again. When the process stopped, all the cash was inside banks and new sight deposits equal to $\frac{\text{additional cash}}{\alpha}$ had been created. Thus M1 had increased, and by considerably more than the additional cash.

When the deposits increased, at the original interest rate, the supply of money exceeded the demand for it and so the demand for bonds exceeded their supply. The price of bonds rose and, by inference, their interest rate decreased. The decrease in interest rate would increase both planned investment and planned consumption. In the former case, the NPVs of additions to plant and machinery would increase so firms wished to invest more. In the latter case, the PV of each consumer's expected future income and their current wealth would increase. Thus each consumer would choose to increase his or her expenditure. If their desired expenditure exceeds their current income, they may reduce their stock of savings or borrow. Therefore, aggregate demand increases and output rises.

Notice that the creation of loans is the result of an increase in the demand for them by firms and households. The increase in demand is due to the fall in interest rates due to the creation of additional deposits by banks.

3.4.3 The credit channel

According to Bernanke and Blinder (1988), the credit channel acts to exaggerate the effect of the money supply on output. There are at least two ways in which the mechanism works: through the balance sheet channel and through the bank lending channel. We consider the balance sheet channel first.

Borrowers (who act as agents) are more aware than lenders (who act as principals) of their ability to repay. To reduce the risk of nonpayment, lenders incur costs in assessing the risks of lending and also of monitoring borrowers. To compensate for the risk, banks may require a higher return than if there was no risk. These costs would not be borne by a firm (or consumer) that used its own funds to finance investment. We call these higher costs moral hazard costs. Therefore, if a firm uses external finance, it would expect to pay a higher interest rate than if it used retained profits. Bernanke and Blinder argue that the difference between these rates, the "external finance premium," is itself positively related to the level of the rates. Thus a change in, say, banks' reserves effects a change in interest rates (as in the interest rate channel) but also in the premium.

In the case of the balance sheet channel, a reduction in banks' reserves makes the financial position of firms worse, the moral hazard costs rise, and so does the premium charged by banks. At the original interest rate, banks supply fewer loans and, provided firms cannot gain loans from others at a lower interest rate than from the banks, firms will invest less than if the channel did not operate. The financial position of firms is made worse in several ways. When banks' reserves fall and the interest rate rises, the interest that firms have to pay on outstanding debt increases. In addition, the value of collateral that a borrower may have offered when he took out the loan decreases. (Think of the value as being the expected present value of the future profits the collateral would generate.) Therefore, the risk for the bank associated with outstanding loans and any new loans would be higher and the interest rate premium charged would be higher. Thus banks would supply fewer loans and firms would plan to invest less. Aggregate demand would fall and so therefore would output.

Similar arguments apply to households' expenditure on durables and housing. For example, when interest rates rise, the interest rate that consumers have to pay on outstanding debt increases and the value of collateral falls. The costs of moral hazard rise, and at the original rate, banks supply fewer loans to consumers. Therefore, the consumers who cannot perfectly substitute loans from other sources reduce their investment expenditure and output decreases.

In the case of the bank lending channel, a reduction in banks' reserves reduces the volume of bank loans directly. This is because, given the reserve ratio α, $\frac{\text{cash reserves}}{\text{deposits}}$, chosen by commercial banks, a reduction in reserves causes banks to reduce deposits. Since α is less than 1, the reduction in reserves is less than the reduction in deposits. Since the total liabilities (that is, deposits) must equal total assets for commercial banks, the banks must reduce the nonreserve elements of their assets also. These nonreserve elements of their assets consist of loans and securities. If banks wish to preserve the ratio of loans to securities (to maintain the same risk-return mix), then they will reduce their supply of loans. The high-risk firms will then have to borrow from nonbank lenders. Banks are assumed to be the most efficient lenders at assessing the riskiness of firms. Therefore, the high-risk firms will have to apply to lenders who will incur more risk assessment costs than banks, and they will have to incur higher search costs than if banks had lent to them. These additional costs increase the financing premium that firms would have to pay and so reduce the NPV of investments that firms wish to make. Thus investment and output would fall. The same process would reduce the supply of loans to consumers for durables and housing purchases.

We now explain Bernanke and Blinder's formal representation of the bank lending channel mechanism. Their model predicts how changes in the money supply or demand for credit, as well as changes in bank reserves, affect output. First, assume that instead of bonds and money being the only assets in an economy, as in the conventional LM model, borrowers and lenders can spread their wealth among bonds, money, and securities. Bernanke and Blinder develop models for the money market (the standard LM model as explained above), the credit market, and the goods market. They assume that the demand for loans depends negatively on their interest rate (they are less expensive) and positively on the return on bonds (bonds are more attractive to buy and can be bought using credit). Demand for credit is also positively related to the number of transactions made to enable purchases, as proxied by output. Now consider the supply of loans. Bernanke and Blinder argue that this depends positively on the loan interest rate because, everything else constant, the return earned by a bank by making a loan is greater, relative to all other uses of its funds, when this rate is higher. The supply is also negatively related to the interest rate on bonds because, holding everything else constant, the lower the rate on bonds, the greater the relative return from making loans. Third, the supply depends positively on the level of deposits above the level banks retain to be able to return cash to customers who have deposits. These relationships are shown in Figure 3.12. Holding everything constant except C_s, C_d, and i_l, the equilibrium will occur at A.

Now consider the market for money. The demand for deposits depends positively on the number of transactions, proxied by output, and negatively on the return on bonds for reasons given earlier. The supply of deposits equals the money multiplier (which is $\frac{1}{\alpha}$; see above) times reserves R. The money multiplier is assumed to depend positively on the return on bonds because the higher this return, the lower the proportion of deposits banks wish to keep in reserve (they prefer to buy bonds) and the multiplier is the inverse of this fraction. The demand and supply of money is represented by Figure 3.13. Using the same explanation as earlier (see section 3.4.1), we can derive the LM curve: combinations of i_b and output at which the demand for deposits equals supply.

Finally, Bernanke and Blinder assume the traditional IS model of the goods market, explained earlier. Thus equilibrium output in this market depends negatively on the interest rate on bonds (a rise in this rate reduces the NPV of investment) and negatively on the interest rate on loans (the greater the interest rate on loans, the fewer loans are demanded and firms invest less and consumer demand decreases also). As aggregate demand decreases, so does output. This relationship is shown in Figure 3.14.

To explain the equilibrium, remember that we require equilibrium in the money market giving the LM curve, which is plotted in Figure 3.15. We require equilibrium in the credit market giving the interest rate on loans. We also require equilibrium in the goods market given the interest rate on loans; this is represented by the line in Figure 3.14, which corresponds to this rate. Call this line XX. Thus conditions represented by LM and XX must hold, so the overall equilibrium levels of i_b and Y must be where the two curves intersect.

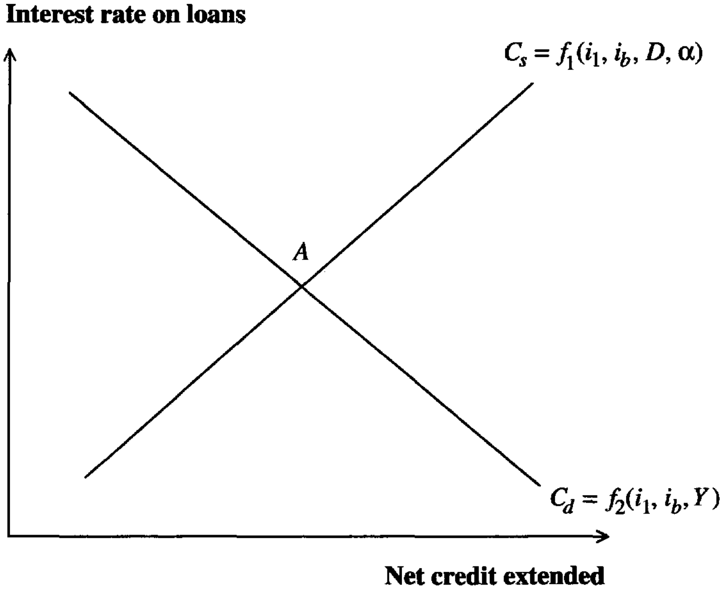


Figure 3.12. *The demand and supply of loans.*

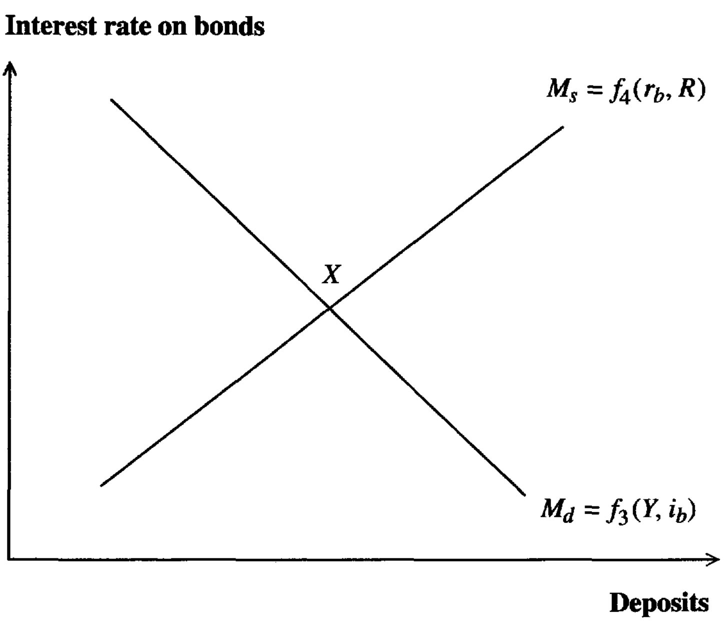


Figure 3.13. *The demand and supply of money.*

This model can be used to show the effects of an increase in the supply of credit at every loan interest rate. This might be the result of banks decreasing their assessment of the level of risk associated with loans. This would shift the supply curve in Figure 3.12 to the right, thus reducing the loan interest rate and increasing the amount of loans. Consumer demand and investment demand would increase and so would output: the XX line in Figure 3.14

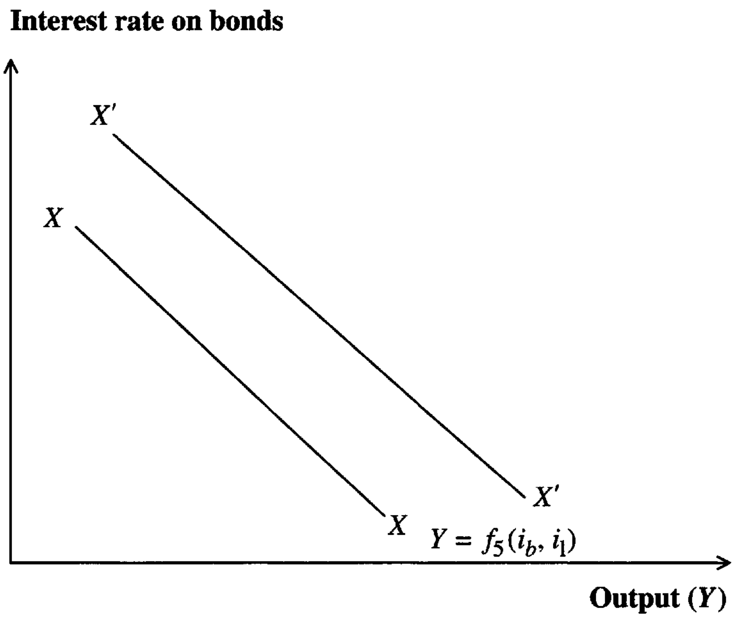


Figure 3.14. *Augmented IS curve.*

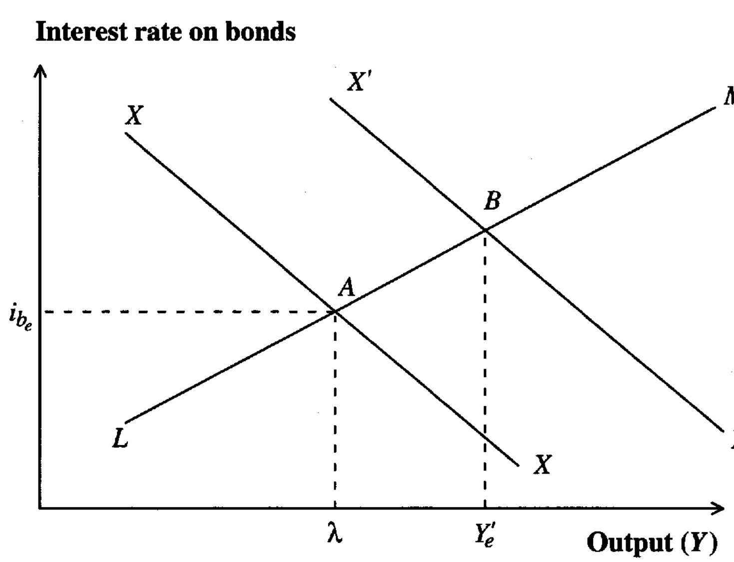


Figure 3.15. *Overall equilibrium.*

shifts to the right to $X'X'$. In Figure 3.15, the new $X'X'$ line is shown, and it can be seen that output would rise from Y_e to Y_e'. Bernanke and Blinder comment that exactly this pattern of adjustments occurred (in reverse) when, in the 1920s, banks believed that loans were more risky. The supply of credit decreased, the loan rate rose, the XX curve moved to the left, and output fell: the Great Depression.

3.4.4 Empirical evidence

A large number of studies have tried to discover whether the credit channel operates as well as the money channel. The conclusions of the studies differ and this is an active area of research. These studies often estimated systems of simultaneous time series regression equations. In each case, a variable was related to lagged values of itself and lagged values of other variables (typically up to six lagged terms each). These are known as vector autoregression (VAR) models and, with some simplification, have the form

$$\mathbf{y}_t = \mathbf{a} + \beta' \mathbf{y}_{t-1} + \mathbf{e}_t,$$

where $\mathbf{y}_t$ is a column vector of variables, β is a matrix of coefficients, and $\mathbf{e}_t$ is a column vector of residuals. Each set of equations allows the researcher to examine the effects of changes in $\mathbf{e}_t$ for one variable on the future values of all of the other variables in $\mathbf{y}_t$.

Using this method, an early study by King (1986) examined the relationships between the growth rates of GNP, deposits, and loans to different sectors of the economy for the U.S. between 1950 and 1979. King found that deposits explained a larger proportion of the variations in future output than did loans, which he interpreted as indicating support for the money channel and little support for the credit channel. However, Bernanke (1986) argued that King imposed ad hoc constraints on the values of the coefficients in the matrix of βs in his equations. When constraints that are economically plausible are applied, both channels appear to be equally important.

The methodology of King and of Bernanke (1986) was criticized by Bernanke and Blinder (1992) as being sensitive to the choice of imposed constraints and specifications of the VARs. Instead, Bernanke and Blinder used a direct measure of monetary policy: the federal funds interest rate. Having estimated the appropriate VARs, they examined the time series paths of bank deposits, bank loans and bank securities, and unemployment to a shock (that is, not explained by other variables in the model) in the funds rate. The funds rate is directly controlled by the Federal Reserve and an increase in it can be seen in Figure 3.10, to correspond to a decrease in the money supply and so a tightening of monetary policy. Bernanke and Blinder's results show that an increase in the funds rate leads to an immediate reduction in both total bank deposits and total bank assets (both values are equal, being on opposite sides of the bank balance sheet). But total bank assets consist of loans and securities, and for the first six months after the shock, it is securities, not loans, that are reduced. After six months, banks' holdings of securities began to rise and they reduce their loans. This continued, and by two years after the shock, banks' holdings of securities had returned to their original levels and loans had decreased by the same amount as deposits.

Similar results were found in other studies. For example, using U.S. data, Kashyap et al. (1993) observed that after each of several specific dates on which monetary contraction occurred, commercial paper held by banks increased and bank loans did not decrease until two years later. Bachetta and Ballabriga (1995), using data for each of 14 European countries, found that following a shock rise in the federal funds rate, the decrease in deposits was initially greater than that in loans (that is, banks initially reduced securities), but after several months, the reverse happened.

However, authors disagree over the interpretation. Bernanke and Blinder (1992) argued that the initial stability of loans occurs because, unlike deposits, they involve a contract and cannot be easily repaid quickly. In the long run, banks can reallocate their portfolio between loans and securities. Therefore, the initial constancy of loans is consistent with a credit channel. Alternatively, Romer and Romer (1990) argued that if a monetary contraction involves banks losing reserves and so demand deposits, banks may respond by issuing deposits (certificates of deposit) whose quantity is not linked to reserves rather than reducing total liabilities. Since total deposits fall hardly at all, neither do total assets and thus loans. Bernanke and Gertler (1995) accepted that this may be valid for the U.S. after 1980, when legal reserve requirements for banks were removed.

A second test of money versus the credit channel is to examine the change in the fraction $\frac{\text{bank loans}}{\text{commercial paper plus bank loans}}$ after monetary tightening. Kayshap et al. (1993) argued that a decrease in this ratio is consistent with a reduction in the supply of debt rather than a reduction in the demand for debt because if there were a decrease in demand, commercial paper and bank loans would decrease by the same proportion and the ratio would remain constant. Kayshap, using U.S. business loan data for 1963–1989, found that the ratio did indeed decrease after monetary tightenings. Ludvigson (1998) found the same results for consumer loans for automobiles between 1965 and 1994.

A third test of the credit channel is to see if there is support for the assumption of the theory that those firms and consumers whose supply of bank debt is reduced are unable to gain debt from other sources. If they can raise debt from nonbank sources, then a reduction in bank debt will not affect aggregate demand and so will not affect output. Kayshap et al. tested this by seeing whether the above ratio significantly explained firms' investment and inventories. They found that it did. A similar result was found by Ludvigson. In addition, Gertler and Gilchrist (1993) argued that small firms were less able to substitute paper for bank credit because of the higher-risk assessment costs to issuers of paper than to banks. Thus the credit channel would predict that an increase in the funds rate would cause small firms to have not only less bank credit but also less of other types of credit and so would cut back investment more than large firms. This is indeed what they found.

Bernanke and Blinder (1992) found that following monetary contraction, eventually unemployment rises at the same time as loans decrease. Romer and Romer (1990) found a consistent result for output. However, this correlation on its own is consistent with both the supply of loans causing a fall in output and with the fall in output reducing the demand for loans: the money channel and the bank lending version of the credit channel. It is also consistent with the balance sheet version of the credit channel, whereby the rise in interest rates makes potential borrowers more risky and so many do not receive credit. Ludvigson (1998) tried to identify which credit channel was operating. He argued that the balance sheet channel implies that given a monetary contraction, riskier borrowers would face a greater reduction in credit supply than less risky borrowers. His data did not support this prediction. Increases in the funds rate led to larger proportionate increases in commercial paper, which is given to riskier borrowers, than in bank credit, which is given to less risky borrowers. Ludvigson argued that, because there is evidence in favor of a credit channel, if the balance sheet channel is not operating, then the bank lending channel must be the mechanism in operation.

A further interesting relationship was found by Dale and Haldane (1995). Using U.K. data for 1974–1992, they found that an increase in interest rates results in an increase in business borrowings in the short term with business lending becoming lower than its original level a year later, after a decrease in output. For the personal sector, a positive shock to interest rates results in an immediate decrease in borrowing, and two years later, the level of deposits becomes less than their original value at the time of the shock. The decrease in lending to the personal sector preceded the decrease in demand, which is consistent with a credit channel mechanism. But the credit channel did not operate for businesses. Furthermore, they found the significance of the credit channel to be small.

3.5 Default behavior

In this section, we consider how default rates vary with the output of an economy. Figure 3.16 shows the GDP (at constant prices) and the default rate on all consumer loans for the U.S. Both series have been detrended so we are observing fluctuations around their long-term trend values. The immediate impression is that the default rate is negatively related to output and that the timing of the peaks (troughs) in output occur, in most cases, in the same quarter as the troughs (peaks) in the default rate.

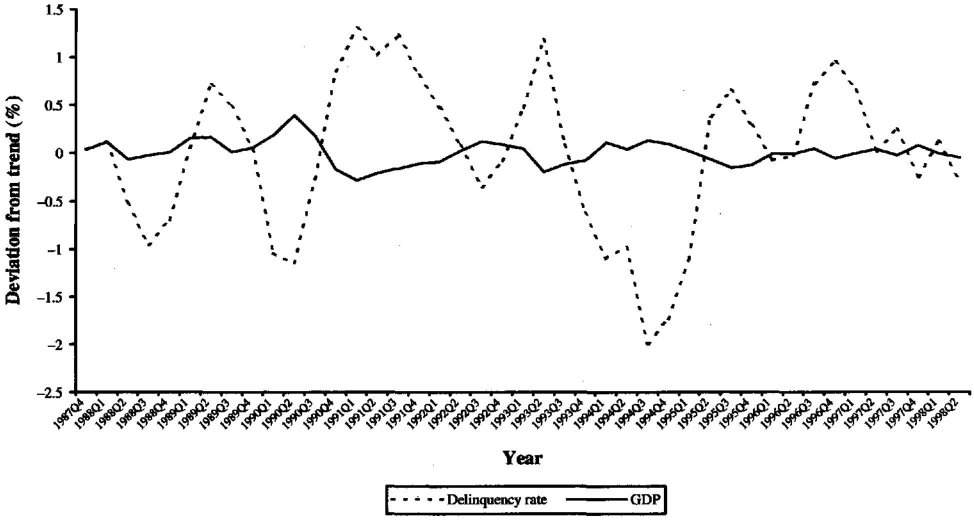


Figure 3.16. *U.S. cyclical behavior of delinquency and business cycles.*

There is no corresponding data for the U.K. The most aggregated data relates to mortgage arrears and is collected by the Council of Mortgage Lenders. Figure 3.17 shows the proportion of mortgages between 6 months and 12 months in arrears and more than 12 months in arrears as well as the number of homes repossessed and GDP. Although the frequency of the data is annual, the same pattern as was observed for the U.S. for all consumer loans is evident.

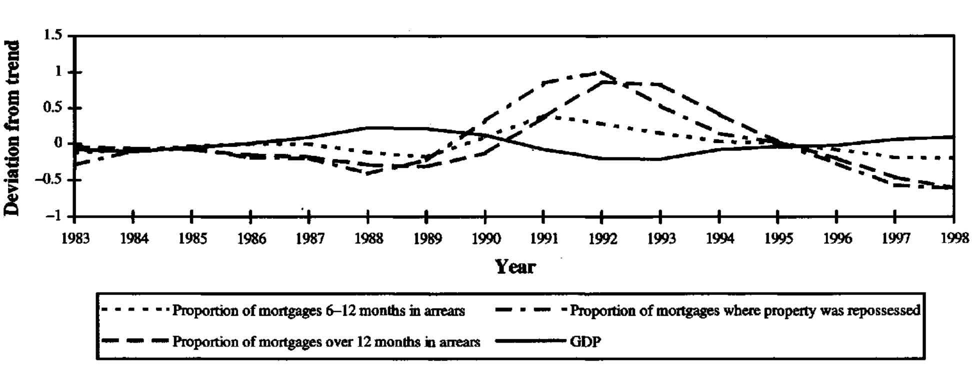


Figure 3.17. *U.K. cyclical behavior of mortgage arrears. (Data from the Council of Mortgage Lenders.)*

Little work has been done to explain patterns of default over time. Sullivan (1987) made the most thorough study and we discuss her results here. Sullivan argued that the default rate varies over time because of changes in the ability of borrowers to repay and because of changes in lenders' credit granting and collection policies. She argued that a borrower will default when the costs of maintaining the loan payments are greater than the costs of not doing so. The costs of the payments are their present value. The costs of not keeping up to date include the repayment costs, λ, which the lender may be able to force onto the borrower, the effect of default on the ability of the borrower to gain credit in the future, and its price. She argued that the costs of maintaining the loan will exceed the costs of default when a borrower's ability to repay deteriorates due to, for example, a decrease in real income. Using aggregate U.S. time series data for the period 1975–1986, she found that the delinquency rate (delinquent for at least 30 days) on consumer installment loans at commercial banks is positively related to total consumer debt outstanding as a proportion of total PDI, and the share that commercial banks had of the consumer loan market, and it is negatively related to the growth rate of this debt. Other indicators of consumers' ability to repay their loans were not significantly related to the delinquency rate. She also found that the factors affecting the default rate varied between types of lenders. For example, for indirect auto loans borrowed from auto finance companies, she found that the delinquency rate was higher in periods of high unemployment and in periods when the debt-to-income ratio was high. She also found that delinquencies increased in periods of high inflation but decreased as the share of the loan market taken by such finance companies increased. Alternatively, for indirect loans borrowed from commercial banks, unemployment and inflation had no effect. The higher the debt-to-income ratio and the market share of banks, the higher the delinquency rate. These differences may reflect the different types of borrower from each sector.

Chapter 4

Statistical Methods for Building Credit Scorecards

4.1 Introduction

When credit scoring was first developed in the 1950s and 1960s, the only methods used were statistical discrimination and classification methods. Even today statistical methods are by far the most common methods for building credit scorecards. Their advantage is that they allow one to use knowledge of the properties of sample estimators and the tools of confidence intervals and hypothesis testing in the credit-scoring context. Thus one is able to comment on the likely discriminating power of the scorecard built and the relative importance of the different characteristics (variables) that make up the scorecard in its discrimination. These statistical techniques then allow one to identify and remove unimportant characteristics and ensure that all the important characteristics remain in the scorecard. Thus one is able to build lean and mean scorecards. One can also use this information when one is looking at what changes might be needed in the questions asked of prospective borrowers.

Although statistical methods were the first to be used to build scoring systems and they remain the most important methods, there have been changes in the methods used during the intervening 40 years. Initially, the methods were based around the discrimination methods suggested by Fisher (1936) for general classification problems. This led to a linear scorecard based on the Fisher linear discriminant function. The assumptions that were needed to ensure that this was the best way to discriminate between good and bad potential customers were extremely restrictive and clearly did not hold in practice, although the scorecards produced were very robust. The Fisher approach could be viewed as a form of linear regression, and this led to an investigation of other forms of regression that had less restrictive assumptions to guarantee their optimality and still led to linear scoring rules. By far the most successful of these is logistic regression, which has taken over from the linear regression–discriminant analysis approach as the most common statistical method. Another approach that has found favor over the last 20 years is the classification tree or recursive partitioning approach. With this, one splits the set of applicants into a number of different subgroups depending on their attributes and then classifies each subgroup as satisfactory or unsatisfactory. Although this does not give a weight to each of the attributes as the linear scorecards does, the result is the same—a method for deciding whether a new applicant will be classified as satisfactory or unsatisfactory. All these methods have been used in practice to devise scorecards for commercial organizations, but there is still plenty of experimentation in using statistical methods. The final statistical method discussed in this chapter is one of these experimental methods—the nonparametric approach based on nearest neighbors.

In this chapter, we review each of these methods and the statistical background that underpins the methods. We begin with discriminant analysis. The next three sections look at how the linear discriminant function is arrived at as a classifier by three different approaches to the problem. The first approach is decision making, where one looks to find the rule that minimizes the expected cost in deciding whether to accept a new customer. The second approach is that which motivated Fisher's original work—namely, seeking to find a function that best differentiates between the two groups of satisfactory (good) and unsatisfactory (bad) customers. The third approach is to consider a regression equation that tries to find the best estimate of the likelihood of a client being good. Each of these approaches arrives at the same linear scorecard method for determining good from bad clients. Sections 4.5 and 4.6 build on the regression approach by looking at regression approaches where the dependent variable is some nonlinear function of the probability of a client being good. One such function leads to logistic regression; other functions lead to probit and tobit analysis. Section 4.7 deals with the classification tree approach and section 4.8 with the nearest-neighbor approach, which are quite different statistical formulations of the original classification problem. Finally, there is a discussion of how some of the methods could be extended to allow for credit-scoring systems where one wants to classify clients into more than the two categories of good and bad.

4.2 Discriminant analysis: Decision theory approach

The credit-granting process leads to a choice between two actions—give this new applicant credit or refuse this applicant credit. Credit scoring tries to assist this decision by finding what would have been the best rule to apply on a sample of previous applicants. The advantage of doing this is that we know how these applicants subsequently performed. If there are only two actions possible—accept or reject—then there is no advantage in classifying this performance into more than two classes—good and bad. Good is any performance that is acceptable to the lending organization, while bad is performance that means the lender wishes he had rejected the applicant. This is looked at more closely in section 8.3. In some organizations, bad performance is taken as missing a given number of consecutive payments, while in others it is the total number of missed payments that matters.

There is an inherent bias in this approach in that the sample is of previous applicants who were granted credit and there is no information on the performance of those applicants who were rejected in the past. Thus the sample is representative of those accepted in the past and is not representative of those who applied in the past. How organizations deal with this bias is looked at in section 8.9.

There is an argument that there are more than the two choices of accept and reject in the credit granting process. For example, one might decide to request further information on the applicant or that this applicant should be considered by a credit analyst for a manual decision. However, these variations are more to do with the ways a lender organizes decision making, and the final decision in every case will be accept or reject. It is not worthwhile to try to classify which applicants will be in a particular group when the decision on whether they are in that group is wholly that of the lender. Thus even in these multiple action processes it is sensible to classify the applicants into only two groups—good and bad—since the final decision will result in one of two actions.

Let $X = (X_1, X_2, \ldots, X_p)$ be the set of p random variables that describe the information available on an applicant for credit both from the application form and through a credit reference bureau check. Of course, nowadays there may not be a physical form, as the details could be captured on a screen via the Internet or they could be written down by a member of the lender staff in a telephone call. We use the words *variable* and *characteristic* interchangeably to describe a typical X_i: the former when we want to emphasize the random nature of this information between applicants and the latter when we want to recall what sort of information it is. The actual value of the variables for a particular applicant is denoted $\mathbf{x} = (x_1, x_2, \dots, x_p)$. In credit-scoring terminology, the different possible values or answers, x_i, to variable X_i are called the attributes of that characteristic. So if a typical characteristic is the applicant's residential status, then its attributes might be owner, rent unfurnished, rent furnished, living with parents, or other. Different lenders may have different groups of attributes of the same characteristic. Thus another lender may decide to classify residential status into owner with no mortgage, owner with mortgage, renting unfurnished property, renting furnished property, subletting property, mobile home, accommodation provided, living with parents, living with other than parents, or other status. It is not uncommon for attribute and characteristic to be mixed up. An easy way to remember which is which is that the Attribute is the Answer to the application form question and the Characteristic is the Cw(qu)estion that was asked.

Returning to the decision that has to be made by the lending organization, suppose A is the set of all possible values that the application variables $\mathbf{X} = (X_1, X_2, \dots, X_p)$ can take, i.e., all the different ways the application form can be answered. The objective is to find a rule that splits the set A into two subsets A_G and A_B so that classifying applicants whose answers are in A_G as "goods" and accepting them while classifying those whose answers are in A_B as "bads" and rejecting them minimizes the expected cost to the lender. The two types of cost correspond to the two types of error that can be made in this decision. One can classify someone who is good as a bad and hence reject the person. In that case the potential profit from that applicant is lost. Assume for now that the expected profit is the same for each applicant and is L. The second error is to classify a bad as a good and so accept the applicant. In that case a debt will be incurred when the customer defaults on the loan. We assume that the expected debt incurred is the same for all customers and is set at D.

Assume that p_G is the proportion of applicants who are goods.

Similarly, let p_B be the proportion of applicants who are bads.

Assume the application characteristics have a finite number of discrete attributes so that A is finite and there are only a finite number of different attributes $\mathbf{x}$. This is like saying there is only a finite number of ways of filling in the application form. Let $p(\mathbf{x}|G)$ be the probability that a good applicant will have attributes $\mathbf{x}$. This is a conditional probability and represents the ratio

$$p(\mathbf{x}|G) = \frac{\text{Prob(applicant is good and has attributes } \mathbf{x})}{\text{Prob(applicant is good)}}. \tag{4.1}$$

Similarly, define $p(x|B)$ to be the probability that a bad applicant will have attributes $\mathbf{x}$.

If $q(G|\mathbf{x})$ is defined to be the probability that someone with application attributes $\mathbf{x}$ is a good, then

$$q(G|\mathbf{x}) = \frac{\text{Prob(applicant has attributes } \mathbf{x} \text{ and is good)}}{\text{Prob(applicant has attributes } \mathbf{x})}, \tag{4.2}$$

and if $p(x) = \text{Prob(applicant has attributes } \mathbf{x})$, then (4.1) and (4.2) can be rearranged to read

$$\text{Prob(applicant has attributes } \mathbf{x} \text{ and is good)} = q(G|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}|G)p_G. \tag{4.3}$$

Hence we arrive at Bayes's theorem, which says

$$q(G|\mathbf{x}) = \frac{p(\mathbf{x}|G)p_G}{p(\mathbf{x})}. \tag{4.4}$$

A similar result holds for $q(B|\mathbf{x})$, the probability that someone with application attributes $\mathbf{x}$ is a bad, namely,

$$q(B|\mathbf{x}) = \frac{p(\mathbf{x}|B)p_B}{p(\mathbf{x})}. \tag{4.5}$$

Note that from (4.4) and (4.5), it follows that

$$\frac{q(G|\mathbf{x})}{q(B|\mathbf{x})} = \frac{p(\mathbf{x}|G)p_G}{p(\mathbf{x}|B)p_B}. \tag{4.6}$$

The expected cost per applicant if we accept applicants with attributes in A_G and reject those with attributes in A_B is

$$L \sum_{\mathbf{x}\in A_B} p(\mathbf{x}|G)p_G + D \sum_{\mathbf{x}\in A_G} p(\mathbf{x}|B)p_B = L \sum_{\mathbf{x}\in A_B} q(G|\mathbf{x})p(\mathbf{x}) + D \sum_{\mathbf{x}\in A_G} q(B|\mathbf{x})p(\mathbf{x}). \tag{4.7}$$

The rule that minimizes this expected cost is straightforward. Consider what the two costs are if we categorize a particular $\mathbf{x} = (x_1, x_2, \dots, x_p)$ into A_G or A_B. If it is put into A_G, then there is only a cost if it is a bad, in which case the expected cost is $Dp(\mathbf{x}|B)p_B$. If $\mathbf{x}$ is classified into A_B, there is only a cost if it is a good, and so the expected cost is $Lp(\mathbf{x}|G)p_G$. Thus one classifies $\mathbf{x}$ into A_G if $Dp(\mathbf{x}|B)p_B \le Lp(\mathbf{x}|G)p_G$. Thus the decision rule that minimizes the expected costs is given by

$$\begin{aligned} A_G &= \{\mathbf{x}|Dp(\mathbf{x}|B)p_B \le Lp(\mathbf{x}|G)p_G\} = \left\{\mathbf{x}|\frac{D}{L} \le \frac{p(\mathbf{x}|G)p_G}{p(\mathbf{x}|B)p_B}\right\} \\ &= \left\{\mathbf{x}|\frac{D}{L} \le \frac{q(G|\mathbf{x})}{q(B|\mathbf{x})}\right\}, \end{aligned} \tag{4.8}$$

where the last expression follows from (4.6).

One criticism of the above criterion is that it depends on the expected costs D and L, which may not be known. So instead of minimizing the expected cost, one could seek to minimize the probability of committing one type of error while keeping the probability of committing the other type of error at an agreed level. In the credit-granting context, the obvious thing to do is to minimize the level of default while keeping the percentage of applicants accepted at some agreed level. The latter requirement is equivalent to keeping the probability of rejecting good applicants at some fixed level.

Suppose one wants the percentage of applicants accepted (the acceptance rate) to be a. Then A_G must satisfy

$$\sum_{\mathbf{x}\in A_G} p(\mathbf{x}|G)p_G + \sum_{\mathbf{x}\in A_G} p(\mathbf{x}|B)p_B = a \tag{4.9}$$

while at the same time minimizing the default rate

$$\sum_{\mathbf{x}\in A_G} p(\mathbf{x}|B)p_B.$$

If we define $b(\mathbf{x}) = p(\mathbf{x}|B)p_B$ for each $x \in A$, then one wants to find the set A_G so that we can

$$\text{minimize} \sum_{\mathbf{x}\in A_G} b(\mathbf{x}) = \sum_{\mathbf{x}\in A_G} \left(\frac{b(\mathbf{x})}{p(\mathbf{x})}\right) p(\mathbf{x}) \quad \text{subject to} \sum_{\mathbf{x}\in A_G} p(\mathbf{x}) = a. \tag{4.10}$$

Using Lagrange multipliers (or common sense using the greedy principle), one can see that this must be the set of attributes $\mathbf{x}$, where $\frac{b(\mathbf{x})}{p(\mathbf{x})} \le c$, where c is chosen so that the sum of the $p(\mathbf{x})$ that satisfy this constraint is equal to a. Hence

$$\begin{aligned} A_G &= \left\{\mathbf{x}|\frac{b(\mathbf{x})}{p(\mathbf{x})} \le c\right\} = \{\mathbf{x}|q(B|\mathbf{x}) \le c\} \\ &= \left\{\mathbf{x}|\frac{1-c}{c} \le \frac{p(\mathbf{x}|G)p_G}{p(\mathbf{x}|B)p_B}\right\}, \end{aligned} \tag{4.11}$$

where the second inequality follows from the definitions of $p(\mathbf{x})$ and $b(\mathbf{x})$.

Thus the form of the decision rule under this criterion is the same as (4.8), the decision rule under the expected cost criterion for some suitable choice of costs D and L.

The whole analysis could be repeated assuming the application characteristics are continuous and not discrete random variables. The only difference would be that the conditional distribution functions $p(\mathbf{x}|G)$, $p(\mathbf{x}|B)$ are replaced by conditional density functions $f(\mathbf{x}|G)$, $f(\mathbf{x}|B)$ and the summations are replaced by integrals. So the expected cost if one splits the set A into sets A_G and A_B and accepts only those in A_G becomes

$$L\int_{\mathbf{x}\in A_B} f(\mathbf{x}|G)p_G d\mathbf{x} + D\int_{\mathbf{x}\in A_G} f(\mathbf{x}|B)p_B d\mathbf{x}, \tag{4.12}$$

and the decision rule that minimizes this is the analogue of (4.8), namely,

$$A_G = \{\mathbf{x}|Df(\mathbf{x}|B)p_B \le Lf(\mathbf{x}|G)p_G\} = \left\{\mathbf{x}|\frac{Dp_B}{Lp_G} \le \frac{f(\mathbf{x}|G)}{f(\mathbf{x}|B)}\right\}. \tag{4.13}$$

4.2.1 Univariate normal case

Consider the simplest possible case where there is only one continous characteristic variable X and its distribution among the goods $f(x|G)$ is normal with mean μ_G and variance σ^2, while the distribution among the bads is normal with mean μ_B and variance σ^2. Then

$$f(x|G) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(\frac{-(x-\mu_G)^2}{2\sigma^2}\right),$$

and so the rule of (4.13) becomes

$$\begin{aligned} \frac{f(x|G)}{f(x|B)} &= \frac{\exp\left(\frac{-(x-\mu_G)^2}{2\sigma^2}\right)}{\exp\left(\frac{-(x-\mu_B)^2}{2\sigma^2}\right)} = \exp\left(\frac{-(x-\mu_G)^2+(x-\mu_B)^2}{2\sigma^2}\right) \ge \frac{Dp_B}{Lp_G} \\ &\Rightarrow x(\mu_G-\mu_B) \ge \frac{\mu_G^2-\mu_B^2}{2} + \sigma^2\log\left(\frac{Dp_B}{Lp_G}\right). \end{aligned} \tag{4.14}$$

Thus the rule becomes "accept if the x value is large enough."

4.2.2 Multivariate normal case with common covariance

A more realistic example is when there are p characteristics (variables) in the application information and the outcomes of these among the goods and the bads both form multivariate normal distributions. Assume the means are $\boldsymbol{\mu}_G$ among the goods and $\boldsymbol{\mu}_B$ among bads with common covariance matrix $\boldsymbol{\Sigma}$. This means that $E(X_i|G) = \mu_{G,i}$, $E(X_i|B) = \mu_{B,i}$, and $E(X_iX_j|G) = E(X_iX_j|B) = \Sigma_{ij}$.

The corresponding density function in this case is

$$f(\mathbf{x}|G) = (2\pi)^{-\frac{p}{2}}(\det\Sigma)^{-\frac{1}{2}}\exp\left(\frac{-(\mathbf{x}-\boldsymbol{\mu}_G)\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_G)^T}{2}\right), \tag{4.15}$$

where $(\mathbf{x}-\boldsymbol{\mu}_G)$ is a vector with 1 row and p columns and $(\mathbf{x}-\boldsymbol{\mu}_G)^{\mathrm{T}}$ is its transpose, which is the same numbers represented as a vector with p rows and 1 column. Following the calculation in (4.14), we get

$$\begin{aligned} &\frac{f(\mathbf{x}|G)}{f(\mathbf{x}|B)} \geq \frac{Dp_B}{Lp_G} \\ \Rightarrow\ &\mathbf{x}\cdot\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_G-\boldsymbol{\mu}_B)^{\mathrm{T}} \geq \frac{\boldsymbol{\mu}_G\cdot\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_G^{\mathrm{T}}-\boldsymbol{\mu}_B\cdot\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_B^{\mathrm{T}}}{2}+\log\left(\frac{Dp_B}{Lp_G}\right). \end{aligned} \tag{4.16}$$

The left-hand side of (4.16) is a weighted sum of the values of the variables, namely, $x_1w_1 + x_2w_2 + \cdots + x_pw_p$, while the right-hand side is a constant. Hence (4.16) leads to a linear scoring rule, which is known as the linear discriminant function.

The above example assumed that the means and covariances of the distributions were known. This is rarely the case, and it is more normal to replace them by the estimates, namely, the sample means $\mathbf{m}_G$ and $\mathbf{m}_B$ and the sample covariance matrix $\mathbf{S}$. The decision rule (4.16) then becomes

$$\mathbf{x}\cdot\mathbf{S}^{-1}(\mathbf{m}_G-\mathbf{m}_B)^T \geq \frac{\mathbf{m}_G\cdot\mathbf{S}^{-1}\mathbf{m}_G^{\mathrm{T}}-\mathbf{m}_B\cdot\mathbf{S}^{-1}\mathbf{m}_B^{\mathrm{T}}}{2}+\log\left(\frac{Dp_B}{Lp_G}\right). \tag{4.17}$$

4.2.3 Multivariate normal case with different covariance matrices

Another obvious restriction on the previous case is that the covariance matrices are the same for the population of goods and of bads. Suppose the covariance matrix on the population of the goods is $\boldsymbol{\Sigma}_G$ for the goods and $\boldsymbol{\Sigma}_B$ for the bads. In this case, (4.16) becomes

$$\begin{aligned} &\frac{f(\mathbf{x}|G)}{f(\mathbf{x}|B)} \geq \frac{Dp_B}{Lp_G} \\ &\Rightarrow \exp\left\{-\frac{1}{2}((\mathbf{x}-\boldsymbol{\mu}_G)\boldsymbol{\Sigma}_G^{-1}(\mathbf{x}-\boldsymbol{\mu}_G)^{\mathrm{T}}-(\mathbf{x}-\boldsymbol{\mu}_B)\boldsymbol{\Sigma}_B^{-1}(\mathbf{x}-\boldsymbol{\mu}_B)^{\mathrm{T}})\right\} \geq \frac{Dp_B}{Lp_G} \\ &\Rightarrow (\mathbf{x}(\boldsymbol{\Sigma}_G^{-1}-\boldsymbol{\Sigma}_B^{-1})\mathbf{x}^{\mathrm{T}}+2\mathbf{x}\cdot(\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\mu}_G^{\mathrm{T}}-\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\mu}_B^{\mathrm{T}}) \geq (\boldsymbol{\mu}_G\boldsymbol{\Sigma}_G^{-1}\boldsymbol{\mu}_G^{\mathrm{T}}+\boldsymbol{\mu}_B\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\mu}_B^{\mathrm{T}}) \\ &\quad +2\log\left(\frac{Dp_B}{Lp_G}\right). \end{aligned} \tag{4.18}$$

The left-hand side here is a quadratic in the values $x_1, x_2, \ldots, x_p$. This appears to be a more general decision rule and so one would expect it to perform better than the linear rule. In practice, however, one has to estimate double the number of parameters $\boldsymbol{\Sigma}_B$ and $\boldsymbol{\Sigma}_G$. The extra uncertainty involved in these estimates makes the quadratic decision rule less robust than the linear one, and in most cases it is not worth trying to get the slightly better accuracy that may come from the quadratic rule. This is confirmed by the work of Reichert, Cho, and G. M. Wagner, who made such comparisons (Reichert et al. 1983).

4.3 Discriminant analysis: Separating the two groups

In Fisher's original work (1936), which introduced the linear discriminant function, the aim was to find the combination of variables that best separated two groups whose characteristics were available. These two groups might be different subspecies of a plant and the characteristics are the physical measurements, or they might be those who survive or succumb to some traumatic injury and the characteristics are the initial responses to various tests. In the credit-scoring context, the two groups are those classified by the lender as goods and bads and the characteristics are the application form details and the credit bureau information.

Let $Y = w_1X_1 + w_2X_2 + \cdots + w_pX_p$ be any linear combination of the characteristics $\mathbf{X} = (X_1, X_2, \ldots, X_p)$. One obvious measure of separation is how different are the mean values of Y for the two different groups of goods and bads in the sample. Thus one looks at the difference between $E(Y|G)$ and $E(Y|B)$ and chooses the weights w_i with $\sum_i w_i = 1$, which maximize this difference. However, this is a little naive because it would say the two groups in Figure 4.1 were equally far apart. What that example shows is that one should also allow for how closely each of the two groups cluster together when one is discussing their separation.

BB B G | B G B G B | G G G

mean of bads mean of goods

B B B | BB B G GG | GG G

mean of bads mean of goods

Figure 4.1. *Two examples where means of goods and bads are the same.*

Fisher suggested that if we assume the two groups have a common sample variance, then a sensible measure of separation is

$$M = \frac{\text{distance between sample means of two groups}}{(\text{sample variance of each group})^{\frac{1}{2}}}.$$

One divides by the square root of the sample variance so as to make the measure scale independent. If one changes the variable from Y to cY, then the measure M does not change.

Assume sample means are $\mathbf{m}_G$ and $\mathbf{m}_B$ for the goods and the bads, respectively, and $\mathbf{S}$ is the common sample variance. If $Y = w_1X_1 + w_2X_2 + \cdots + w_pX_p$, then the corresponding separating distance M would be

$$M = \mathbf{w}^{\mathrm{T}} \cdot \frac{\mathbf{m}_G - \mathbf{m}_B}{(\mathbf{w}^{\mathrm{T}} \cdot \mathbf{S} \cdot \mathbf{w})^{\frac{1}{2}}}. \tag{4.19}$$

This follows since $E(Y|G) = \mathbf{w} \cdot \mathbf{m}_G^{\mathrm{T}}$, $E(Y|B) = \mathbf{w} \cdot \mathbf{m}_B^{\mathrm{T}}$, and $\mathrm{Var}(Y) = \mathbf{w} \cdot \mathbf{S} \cdot \mathbf{w}^{\mathrm{T}}$. Differentiating this with respect to $\mathbf{w}$ and setting the derivative equal to 0 shows that this value M is maximized when

$$\frac{\mathbf{m}_G - \mathbf{m}_B}{(\mathbf{w} \cdot \mathbf{S} \cdot \mathbf{w}^{\mathrm{T}})^{\frac{1}{2}}} - \frac{(\mathbf{w} \cdot (\mathbf{m}_G - \mathbf{m}_B)^{\mathrm{T}})(\mathbf{S}\mathbf{w}^{\mathrm{T}})}{(\mathbf{w} \cdot \mathbf{S} \cdot \mathbf{w}^{\mathrm{T}})^{\frac{3}{2}}} = 0,$$
$$(\mathbf{m}_G - \mathbf{m}_B)(\mathbf{w} \cdot \mathbf{S} \cdot \mathbf{w}^{\mathrm{T}}) = (\mathbf{S}\mathbf{w}^{\mathrm{T}})(\mathbf{w} \cdot (\mathbf{m}_G - \mathbf{m}_B)^{\mathrm{T}}). \tag{4.20}$$

In fact, all this shows is that it is a turning point, but the fact that the second derivatives of M with respect to $\mathbf{w}$ form a positive definite matrix guarantees that it is a minimum. Since $\frac{\mathbf{w} \cdot \mathbf{S} \cdot \mathbf{w}^{\mathrm{T}}}{(\mathbf{w} \cdot (\mathbf{m}_G - \mathbf{m}_B)^{\mathrm{T}})}$ is a scalar λ, this gives

$$\mathbf{w}^T \propto (\mathbf{S}^{-1}(\mathbf{m}_G - \mathbf{m}_B)^{\mathrm{T}}). \tag{4.21}$$

Thus the weights are the same as those obtained in (4.17), although there has been no assumption of normality this time. It is just the best separator of the goods and the bads under this criterion no matter what their distribution. This result holds for all distributions because the distance measure M involves only the mean and variance of the distributions and thus gives the same results for all distributions with the same mean and variance.

Figure 4.2 shows graphically what the scorecard (4.21) seeks to do. The $\mathbf{S}^{-1}$ term standardizes the two groups so they have the same dispersion in all directions. $\mathbf{w}$ is then the direction joining the means of the goods and the bads after they have been standardized. Thus a line perpendicular to this line joins the two means. The cutoff score is then the midpoint of the distance between the means of the standardized groups.

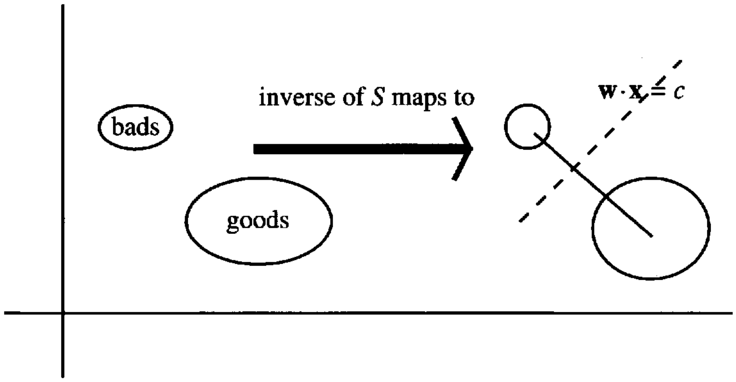


Figure 4.2. *Line corresponding to scorecard.*

4.4 Discriminant analysis: A form of linear regression

Another approach to credit scoring, which again arrives at the linear discriminant function, is linear regression. In this, one tries to find the best linear combination of the characteristics

$$w_0 + w_1X_1 + w_2X_2 + \cdots + w_pX_p = \mathbf{w}^* \cdot \mathbf{X}^{*^{\mathrm{T}}},$$

where

$$\mathbf{w}^* = (w_0, w_1, w_2, \ldots, w_p), \qquad \mathbf{X}^* = (1, X_1, X_2, \ldots, X_p),$$

which explains the probability of default. So if p_i is the probability that applicant i in the sample has defaulted, one wants to find $\mathbf{w}^*$ to best approximate

$$p_i = w_0 + x_{i1}w_1 + x_{i2}w_2 + \cdots + x_{ip}w_p \quad \text{for all } i. \tag{4.22}$$

Suppose n_G of the sample are goods; then for ease of notation, we assume these are the first n_G in the sample and so $p_i = 1$ for $i = 1, \ldots, n_G$. The remaining n_B of the sample $i = n_G + 1, \ldots, n_G + n_B$ are bad, so for them $p_i = 0$, where $n_G + n_B = n$.

In linear regression, we choose the coefficient that minimizes the mean square error between the left- and right-hand sides of (4.22). This corresponds to minimizing

$$\sum_{i=1}^{n_G}\left(1 - \sum_{j=0}^{p} w_j x_{ij}\right)^2 + \sum_{i=n_G+1}^{n_G+n_B}\left(\sum_{j=0}^{p} w_j x_{ij}\right)^2. \tag{4.23}$$

In vector notation, (4.22) can be rewritten as

$$\begin{pmatrix} 1 & \mathbf{X}_G \\ 1 & \mathbf{X}_B \end{pmatrix}\begin{pmatrix} w_0 \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{1}_G \\ \mathbf{0} \end{pmatrix} \quad \text{or} \quad \mathbf{Y}\mathbf{w}^{\mathrm{T}} = \mathbf{b}^{\mathrm{T}}, \tag{4.24}$$

where

$$\mathbf{Y} = \begin{pmatrix} \mathbf{1}_G & \mathbf{X}_G \\ \mathbf{1}_B & \mathbf{X}_B \end{pmatrix}$$

is an n-row $\times$ $(p+1)$-column matrix,

$$\mathbf{X}_G = \begin{pmatrix} x_{11} & \cdots & \cdots & x_{1p} \\ x_{21} & \cdots & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n_G 1} & \cdots & \cdots & x_{n_G p} \end{pmatrix}$$

is an $n_G \times p$ matrix,

$$\mathbf{X}_B = \begin{pmatrix} x_{n_G+1 1} & \cdots & \cdots & x_{n_G+1 p} \\ \vdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ x_{n_G+n_B 1} & \cdots & \cdots & x_{n_G+n_B p} \end{pmatrix}$$

is an $n_B \times p$ matrix, and

$$\mathbf{b}^{\mathrm{T}} = \begin{pmatrix} \mathbf{1}_G \\ 0 \end{pmatrix},$$

where $\mathbf{1}_G$ ($\mathbf{1}_B$) is the $1 \times n_G$ (n_B) vector with all entries 1.

Finding the coefficients of the linear regression corresponds as in (4.23) to

$$\text{Minimize } (\mathbf{Y}\mathbf{w}^{\mathrm{T}} - \mathbf{b}^{\mathrm{T}})^{\mathrm{T}}(\mathbf{Y}\mathbf{w}^{\mathrm{T}} - \mathbf{b}^{\mathrm{T}}). \tag{4.25}$$

Differentiating with respect to $\mathbf{w}$ says this is minimized when the derivative is zero; i.e.,

$$\mathbf{Y}^{\mathrm{T}}(\mathbf{Y}\mathbf{w}^{\mathrm{T}} - \mathbf{b}^{\mathrm{T}}) = \mathbf{0} \quad \text{or} \quad \mathbf{Y}^{\mathrm{T}}\mathbf{Y}\mathbf{w}^{\mathrm{T}} = \mathbf{Y}^{\mathrm{T}}\mathbf{b}^{\mathrm{T}},$$

$$\mathbf{Y}^{\mathrm{T}} \cdot \mathbf{b}^{\mathrm{T}} = \begin{pmatrix} 1 & 1 \\ \mathbf{X}_G & \mathbf{X}_B \end{pmatrix} \cdot \begin{pmatrix} \mathbf{1}_G \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} n_G \\ n_G \mathbf{m}_G \end{pmatrix},$$

$$\text{and} \quad \mathbf{Y}^{\mathrm{T}}\mathbf{Y} = \begin{pmatrix} 1 & 1 \\ \mathbf{X}_G & \mathbf{X}_B \end{pmatrix}\begin{pmatrix} 1 & \mathbf{X}_G \\ 1 & \mathbf{X}_B \end{pmatrix} = \begin{pmatrix} n & n_G\mathbf{m}_G + n_B\mathbf{m}_B \\ n_G\mathbf{m}_G^{\mathrm{T}} + n_B\mathbf{m}_B^{\mathrm{T}} & \mathbf{X}_G^{\mathrm{T}}\mathbf{X}_G + \mathbf{X}_B^{\mathrm{T}}\mathbf{X}_B \end{pmatrix}. \tag{4.26}$$

If for explanatory purposes we denote sample expectations as actual expectations, we get

$$\mathbf{X}_G^{\mathrm{T}}\mathbf{X}_G + \mathbf{X}_B^{\mathrm{T}}\mathbf{X}_B = nE\{X_iX_j\} = n\,\mathrm{Cov}(X_i, X_j) + n_G\mathbf{m}_G\mathbf{m}_G^{\mathrm{T}} + n_B\mathbf{m}_B\mathbf{m}_B^{\mathrm{T}}.$$

If S is the sample covariance matrix, this gives

$$\mathbf{X}_G^{\mathrm{T}}\mathbf{X}_G + \mathbf{X}_B^{\mathrm{T}}\mathbf{X}_B = n\mathbf{S} + n_G\mathbf{m}_G\mathbf{m}_G^{\mathrm{T}} + n_B\mathbf{m}_B\mathbf{m}_B^{\mathrm{T}}. \tag{4.27}$$

Expanding (4.26) and using (4.27) gives

$$\begin{aligned} nw_0 + (n_G\mathbf{m}_G + n_B\mathbf{m}_B)\mathbf{w}^{\mathrm{T}} &= n_G, \\ (n_G\mathbf{m}_G^{\mathrm{T}} + n_B\mathbf{m}_B^{\mathrm{T}})w_0 + (n\mathbf{S} + n_G\mathbf{m}_G\mathbf{m}_G^{\mathrm{T}} + n_B\mathbf{m}_B\mathbf{m}_B^{\mathrm{T}})\mathbf{w}^{\mathrm{T}} &= n_G\mathbf{m}_G^{\mathrm{T}}. \end{aligned} \tag{4.28}$$

Substituting the first equation in (4.28) into the second one gives

$$\begin{aligned} &((n_G\mathbf{m}_G^{\mathrm{T}} + n_B\mathbf{m}_B^{\mathrm{T}})(n_G - (n_G\mathbf{m}_G + n_B\mathbf{m}_B)\mathbf{w}^{\mathrm{T}})/n) \\ &\qquad + (n_G\mathbf{m}_G\mathbf{m}_G^{\mathrm{T}} + n_B\mathbf{m}_B\mathbf{m}_B^{\mathrm{T}})\mathbf{w}^{\mathrm{T}} + n\mathbf{S}\mathbf{w}^{\mathrm{T}} = n_G\mathbf{m}_G^{\mathrm{T}}, \\ &\qquad \text{so } \left(\frac{n_Gn_B}{n}\right)(\mathbf{m}_G - \mathbf{m}_B)\mathbf{w}^{\mathrm{T}} + n\mathbf{S}\mathbf{w}^{\mathrm{T}} = \left(\frac{n_Gn_B}{n}\right)(\mathbf{m}_G - \mathbf{m}_B)^{\mathrm{T}}; \\ &\qquad\qquad \text{thus } \mathbf{S}\mathbf{w}^{\mathrm{T}} = c(\mathbf{m}_G - \mathbf{m}_B)^{\mathrm{T}}. \end{aligned} \tag{4.29}$$

Thus (4.29) gives the best choice of $\mathbf{w} = (w_1, w_2, \ldots, w_p)$ for the coefficients of the linear regression. This is the same $\mathbf{w}$ as in (4.21), namely, the linear discriminant function. This approach shows, however, that one can obtain the coeffieients of the credit scorecard by the least-squares approach beloved of linear regression.

We have taken the obvious left-hand sides in the regression equation (4.22), where goods have a value 1 and bads have a value 0. These have given a set of constants, which we label $\mathbf{w}(\mathbf{1}, \mathbf{0})^*$. If one takes any other values so that the goods have a left-hand side of g and the bads have a left-hand side of b, then the coefficients in the regression—$\mathbf{w}(\mathbf{g}, \mathbf{b})^*$—differ only in the constant term w_0 since

$$\mathbf{w}(\mathbf{a}, \mathbf{b})^* = b + (g - b)\mathbf{w}(\mathbf{1}, \mathbf{0})^*. \tag{4.30}$$

4.5 Logistic regression

The regression approach to linear discrimination has one obvious flaw. In (4.22), the right-hand side could take any value from $-\infty$ to $+\infty$, but the left-hand side is a probability and so should take only values between 0 and 1. It would be better if the left-hand side were a function of p_i, which could take a wider range of values. Then one would not have the difficulty that all the data points have very similar values of the dependent variables or that the regression equation predicts probabilities that are less than 0 or greater than 1. One such function is the log of the probability odds. This leads to the logistic regression approach, on which Wiginton (1980) was one of the first to publish credit-scoring results. In logistic regression, one matches the log of the probability odds by a linear combination of the characteristic variables; i.e.,

$$\log\left(\frac{p_i}{1 - p_i}\right) = w_0 + w_1x_1 + w_2x_2 + \cdots + w_px_p = \mathbf{w} \cdot \mathbf{x}^{\mathrm{T}}. \tag{4.31}$$