ANGLIA RUSKIN UNIVERSITY

BRITISH BANKING-HALLS
AS A PROPERTY INVESTMENT

MALVERN TIPPING

A Thesis in partial fulfilment of the
requirements of Anglia Ruskin University
for the degree of Professional Doctorate

Submitted July 2011

ACKNOWLEDGEMENTS

I have been fortunate to have had professional colleagues and work colleagues who have not only encouraged my research, but who have positively supported it. In particular, I thank the late Dr. Richard K. Bullard, who initiated the Built Environment Professional Doctorate programme at the university and who impressed upon me the need to become a part of that programme from the outset. I am especially indebted to his guidance on how to become published.

Acknowledgements are also given to the late William Knight, who both as a fellow candidate and as the East of England regional vice-chairman of the Royal Institution of Chartered Surveyors, gave equal encouragement and advice. Use of his private library in Norwich was especially helpful.

The dedication and incisive guidance of my principal supervisor, Dr. Terence Lam, has been much appreciated. At times a hard task-master, he has striven to ensure that my research has evolved into a study that is fully robust. He, too, has equipped me with skills that will enable me to research and publish other work in the future.

Ian Frame, as the current course director, has instilled the programme with much enthusiasm. It is that enthusiasm that has carried all candidates forward during both the highs and lows. Both Ian Frame and Peter Wynn are thanked for their respective roles as second and third supervisors.

Roger Newton is thanked for mentoring me in more advanced statistical techniques over the last three years. His tuition, guidance and advice have been invaluable in equipping me with the skills required to undertake the complex, but necessary, statistical analyses, model-building and testing. So much work was undertaken under his guidance, that the final text merely comprises the polished work after the frustration of venturing down several blind alleys. The computer scripts to enable the statistical analyses were written by Roger.

Even in the era of digital technology, librarians have such an important role in assisting the research community. Special thanks are given to Diane Hilton of the university's library at the Rivermead campus for her much appreciated assistance. Also acknowledged are the librarians at both the British Library and the Royal Institution of Chartered Surveyors.

Richard Auterac, Duncan Moir and John Townsend are thanked for the time and trouble that they gave in interviews and discussions with the candidate as leading experts in the field. Those professional members of their staff who were respondents to the questionnaire are similarly acknowledged.

Shabina Bashir of the Investment Property Databank is thanked for the provision of data with respect to the IPD United Kingdom retail property index free of charge.

I thank two of the companies which employ me, Tipping Estates Limited and Devitt & Sons Limited, for their sponsorship and for the use of their resources. Finally, I thank my proof-reader, Anne Crosland, for her patience and endurance.

ANGLIA RUSKIN UNIVERSITY
ABSTRACT

FACULTY OF SCIENCE AND TECHNOLOGY

PROFESSIONAL DOCTORATE

BRITISH BANKING-HALLS
AS A PROPERTY INVESTMENT
By MALVERN TIPPING
July 2011

ABSTRACT

This research is related to British banking-halls as a class of real estate investment. Sale-and-leaseback has become an increasingly common approach during the last two decades for the holding of British banking-halls. One measure used in making property investment decisions is the all risks yield (ARY). Investors and their advisors have a need for a predictive framework which they can use for predicting those retail bank premises likely to achieve the highest ARY when assembling investment portfolios of such properties. A predictive framework necessitates the identification of those factors significantly influencing the yields of British banking-halls. This research aims to develop such a framework.

Triangulation methodology was adopted to establish and test the predictive framework. A literature review established theory before a qualitative study, based upon semi-structured interviews and a questionnaire, was used to establish the influencing factors. A cross-sectional study of auction data then formed the basis of the quantitative regression study.

The qualitative and quantitative studies validated that four factors were significant in influencing yield. These were tenant banking company, lot size, super-region and the macro-economic cycle index.

A toolkit comprising a predictive framework for those banking-halls likely to produce the highest ARY was produced. This is capable of being used by professional practitioners and investors in predicting high yield for portfolio building purposes. The predictive framework was developed based upon the quantitative data from those three banks with the most premises sold by sale-and-leaseback. It formed a baseline from which further studies can build to test its significance for other banks. Consequently, a more robust predictive framework can be developed for banking-hall investments. Further research can also be conducted to develop predictive frameworks forecasting yields for investment in other commercial retail sectors, based upon the findings of this research.

Key words: Banking-hall; investment; portfolio; predictive framework; yield

Table of Contents

List of Figures

List of Tables

List of Appendices

COPYRIGHT

1. Introduction

1.1 The significance

During the past two decades, sale-and-leaseback of property has become a model increasingly adopted by the United Kingdom's largest companies. Much of the wealth of the United Kingdom's largest companies has traditionally been held in real estate. Prevailing theory has been that these companies should release capital from their real estate holdings.

Many of the de-nationalized state industries have released vast amounts of capital through the sale of surplus property. Since 1987 many of the United Kingdom's leading trading companies have released value from their operational property through the process of sale-and-leaseback. This was an approach adopted by the now defunct Woolworths Group plc and its predecessor companies ("Woolworths") in 1987 (Townsend). The sale-and-leaseback model has become much more widespread since then.

The British retail banks have been one of the pre-eminent types of company adopting the sale-and-leaseback approach with respect to their operational properties. In particular, the approach has found favour amongst the leading British retail banks. The latter are defined as Barclays, Lloyds TSB, Royal Bank of Scotland Group (including National Westminster), HSBC and HBOS (Halifax Bank of Scotland). The definition is determined by two separate tests. The first test comprises the five retail banking groups with the largest number of retail branches in Great Britain. The number of branches in each respective retail banking network for 2004, which falls in the middle of the study period, is listed in Table 1.1. The second test comprises the five retail banking groups registered in Great Britain with the largest worldwide assets. Each of the two tests comprises the same five retail banking groups (British Bankers' Association, 2005a and 2005b).

In 1992 Barclays Bank was the first of the principal British retail banks to adopt the model. Since then the model has also been adopted widely by the other principal retail banks: Midland Bank and its successor HSBC Bank, and Lloyds TSB Bank. All of the principal retail banks have had experience of sale-and-leaseback with respect to their operational property to a lesser or a greater extent. Also, several of the secondary retail banks have more recently adopted the model. Having sold several of its operational properties on a sale-and-leaseback basis between 1992 and 1994, Barclays Bank has subjected many more of its operational properties to the model since 2005.

Table 1.1: Main branded bank branch networks in Great Britain in 2004

Bank Name	Branches
Royal Bank of Scotland Group (including National Westminster)	2,273
Barclays (including Woolwich)	2,059
Lloyds TSB (including Lloyds TSB Scotland)	1,976
HSBC	1,569
HBOS (Halifax & Bank of Scotland)	1,064
Abbey	724
Alliance & Leicester	254
Clydesdale Bank	232
Yorkshire Bank	217
Bradford & Bingley	208
Cheltenham & Gloucester	205
Bristol & West	97
The Co-operative Bank	91
Northern Rock	56
Total	**11,025**

Source: British Bankers' Association (2005)

1.2 The knowledge

1.2.1 The existing knowledge

The sale-and-leaseback of British banking-halls has been conducted by a small number of property consultants based in London. Where these properties have been sold as individual lots rather than as parts of portfolios, they have been sold by public auction in London.

The high profile sale-and-leaseback transactions offered at auction in London have attracted keen interest and high demand from private investors. This has been reflected by the low yields at which these properties have sold relative to the overall property investment market at any given time. Ninety-seven Midland Bank lots sold at an auction in 1997 achieved a mean yield of 7.94 per cent. By 2006, many banking-halls lots were achieving yields below four per cent. Between 1997 and 2006, there has been yield compression in the wider property investment market. However, throughout that period, banking-halls have consistently achieved lower yields than some other property classes at any given time. Examples of property types that maintained higher yields during the period are public houses and convenience stores, both of which were recognized as having weaker covenants.

British banking-halls have always attracted high demand from investors when sold on sale-and-leaseback terms by the retail banks. The freehold interests of many of the banking-halls sold on sale-and-leaseback terms have since returned to the auction market. On re-sale, these freeholds have attracted equally high demand from investors.

The sponsors of this research are interested in devising a framework for the selection of an investment portfolio of freehold banking-halls based on premises giving the highest All Risks Yields (ARYs). Accordingly, the primary purpose of this research is to provide such a framework. It seeks to fill gaps in the knowledge to explain the influences in banking-hall ARYs during the study period with a view to enabling the creation of a predictive framework.

1.2.2 The gap in the knowledge

Some literature exists on the future occupational property needs of the British retail banks. It gives some indication as to possible future real estate needs of the retail banks. In particular, design of property is considered in the limited extant literature.

However, no extant published research on the performance of British banking-halls as a medium of property investment could be identified. Electronic searches have failed to find extant papers on the subject. Moreover, an electronic search of the Index to Theses, formerly the ASLIB Index to Theses, failed to show any thesis as having been published in Britain or Ireland on the subject of British banking-halls as a medium of property investment. Therefore, a gap in the knowledge clearly exists. Findings derived from such research would equip investors and others with the knowledge that would better enable them to participate in the investment process.

The gap in the knowledge on banking-halls as a property investment includes:

- The relevance of the extent of the reversion on let retail banking premises (Cavanagh, 2003 and Maxwell, 2006a)
- The influence of regional location on investment yields (Cavanagh, 2003 and Maxwell, 2006e)
- The influence of lot size on investment yields (Cavanagh, 2003 and Maxwell, 2006e)
- The influence, if any, of the influx of amateur investors on investment yields (Cavanagh, 2003)
- The effects of branch crossover (Property Week, 2007a)

Each of these questions arises from the citations and other sources from within professional practice. In some instances, the influence of the respective question is alluded to. In other instances, the question is directly asked. In neither case is the point proven. Therefore, the gap in the knowledge remains until empirical evidence is collected and tested.

Since most of the British banking-hall sale-and-leaseback investments have been sold at public auction, a great deal of empirical data is extant. Moreover, these data are in the public domain. This renders the data capable of collection, recording and testing. Therefore, several of these gaps in the knowledge are capable of being filled through the collection and analysis of empirical data within the context of the professional practice.

Limits on time and other resources dictate that not all of these gaps in the knowledge are capable of resolution within the context of the present research framework through the analysis of empirical data. Where this is the case, either extant literature is reviewed in order to provide some form of answer or unresolved gaps in the knowledge are recommended for further investigation.

1.3 Defining the title and the scope of research

Giving definition to the title of the area under investigation renders precision to the research. Many words have more than one meaning. Some have several meanings. Therefore, defining each word in the title gives clarity and precision to the area of investigation.

For the purpose of this investigation, the word 'British' is applied to define a geographical and political area. Accordingly, only those banking-halls within the geographical entity of Great Britain are included in this research. Banking-halls occupied by British retail banks, but beyond the boundaries of Great Britain, are not included. Great Britain is defined as the largest island in the British Isles and its outlying islands. It includes the geographical and political entities of England, Scotland and Wales. Great Britain is a component of the United Kingdom, but the United Kingdom is wider than just Great Britain. The United Kingdom comprises Great Britain and Northern Ireland.

The Isle of Man is neither part of the Great Britain nor part of the United Kingdom. It is a British Crown Colony. As self-governing Crown dependencies, the Channel Islands are treated as not being British for the purposes of this study. The Republic of Ireland, although geographically within the British Isles, is an independent and sovereign state, and does not form part of either Great Britain or the United Kingdom.

The term *banking-hall* is commonly accepted by both those in banking and property circles as being the built area within which retail banks transact their everyday business with their retail clientele. English dictionaries give several definitions to the word 'hall'. Together, several of these definitions encapsulate the notion of a large room or space under a roof below which meetings and receptions are held. A banking-hall is a large public room or space under a roof below which banking transactions are conducted. The type of banking normally conducted within a banking-hall is that of a retail nature transacted with members of the general public.

For the purpose of this study, a banking-hall is defined as the dedicated space within a building in which retail banking is transacted together with any ancillary offices. This study treats self-contained premises comprising these elements as banking-halls.

The word 'property' derives from the Latin *proprietas*, which means ownership. In English, property is taken to be something capable of ownership. Property can take many forms. Both land and the buildings upon it are a form of property. Rights attached to land and buildings are also a form of property. Movable goods, such as furniture and livestock, are forms of property. In recent years, greater importance has been attached to what is termed as intellectual property, which includes patents and copyrights.

Concepts of property may vary in different legal jurisdictions. However, most legal jurisdictions distinguish real property from other forms of property. In generic terms, real property is taken to comprise interests in land and the buildings on it, together with certain rights attached to land. Often real property is described as real estate. The legal definition of real property in England and Wales follows that of English Law, which is based upon a common law system. In Scotland, where the law is based upon a Roman system, the legal technicalities do differ from the rest of Great Britain (Paisley, 2000). However, for the purpose of this study, the term 'property' is taken to comprise all real property in the generic sense irrespective of whether the property lies in England, Scotland or Wales.

In the broader sense, investment is normally considered to be the application of something to obtain a profit. Labour might be applied to obtain a profit. However, investment is more generally taken to mean the application of money for gain. This study treats investment as meaning the placing of money into something, such as property, with a view to obtaining a gain, whether actual money or money's worth.

The title of this study, *British Banking-halls as a Property Investment*, encapsulates the focus of this research. It means that the focus of the research is into the potential for economic gain from the purchasing of interests in retail banking premises situated in Great Britain.

1.4 The research question

The principal purpose of this thesis is to determine what is likely to be the most profitable approach for the property investor to invest in freehold British banking-halls. For clarification, the most profitable approach has been defined by the sponsors as being the highest initial yield. The potential for capital growth and rental growth is beyond the scope of this research. Since this research is being conducted within the context of professional practice, the focus of the research is directed towards the data available within professional practice. As far as banking-halls are concerned, the empirical data derive from the process of investment properties being sold at auction. The process readily places in the public domain data with respect to rents, sale prices, yields and regions within which the properties are located.

Obvious areas of investigation arising from the freely available data are rents, sale prices, yields and regions. To this list should be added lot size, which can easily be ascertained from the freely available data. Andrews (2003) suggests that hypotheses may be formulated from 'inductive thinking' before being subjected to empirical thinking. Inductive thinking following review of the literature and drawing on the experience of professional practice might suggest other variables that should be tested as:

- The less reversionary banking-halls are in especially high demand from investors;

- Premises let to the smaller and demutualized building societies do not necessarily command higher yields than those let to the principal retail banks;

- Branch crossover has resulted in closures through rationalization;

- The influx of new, amateur investors has pushed up capital values and, conversely, compressed yields.

Data with respect to branch crossover are not freely available within every day professional practice. Furthermore, limits on time and other resources do not permit the collection of crossover data in this investigation. Therefore, this has been placed outside the conceptual framework.

Similarly, the data freely available from the auction results do not facilitate the testing of the impact of the influx of new, amateur investors. Again, this has been placed outside the conceptual framework.

Since no auction data exists with respect to banking-halls in Northern Ireland, that part of the United Kingdom has been omitted from the investigation.

A significant minority of the banking-halls sold at auction formed only a part of larger premises. Often these larger premises also comprised residential property on floors above banking-halls situated on the ground floor. In other instances, other commercial concerns are located in the same property. In either case, scientific testing of such premises relative both to each other and premises comprising only banking-halls is not possible using the available empirical data. Therefore, premises that are only partly banking-halls have been omitted from the study.

The principal purpose of this research is to identify how returns from investing in freehold British banking-halls might be maximized for the investor. Specifically, the sponsors have stated that they require a toolkit for the prediction of those retail bank premises likely to return a higher initial yield for portfolio building purposes. Therefore, the research is undertaken from the perspective of investors and their professional advisors. Based upon the principal purpose of this thesis, the specific research question is:

"How can property investors select freehold British banking-halls that are likely to provide the highest initial yield on their investment?"

The aims and objectives of this research are to develop a predictive framework for forecasting the yields of banking-halls for investment portfolio purposes. The objectives are:

1. To explore the occupational models of retail bank operators and especially the use of sale-and-leaseback.

2. To identify investment opportunities in banking-halls.

3. Identify the influencing factors with particular reference to:

    - Regional disparity in yields.

- Lot size has an effect inasmuch as there is increased demand for smaller lots.

- Tenant banking company having an effect.

- Time has an effect.

- The appropriate Investment Property Databank (IPD) index has an effect.

4. To develop a statistical model for prediction.

1.5 The contribution of the researcher's Professional Doctorate Study

Professional doctorate study possesses some unique attributes related to professional practice. The present study into the prediction of yields for banking hall investment is therefore intended to achieve the following:

- Contribution to the professional advice activity for banking-hall investment clients in the researcher's company.

- Contribution to the researcher's own professional development in property investment knowledge.

- Contribution to the surveying profession by providing new, specialist skills relating to banking-hall investment for the retail property market.

1.6 Assessment of the sources

This research relies upon more than one source of evidence. The testing of the hypotheses relies upon deductive research. The development and the refinement of the theory rely upon inductive research. The deductive research relies upon the empirical data of banking-hall lots sold at auction. Prior to that, the inductive research is heavily dependent upon the review of extant literature. A diverse range of literature is reviewed during the inductive research process.

The empirical data with respect to the banking-hall premises is derived from lots sold at public auction. These lots have been sold at public auctions conducted by the principal London property auction houses. All of the empirical data is in the public domain. In particular, data on rent reserved, region, tenure and use are published in auction catalogues prior to sale. Not only are these catalogues published in traditional printed format, but also in electronic format. The electronic format is capable of being downloaded via the internet and stored on computer. After each auction, the principal London property auction houses publish the results of their sales. All the data on hammer prices are accordingly published. It is all placed on electronic databases accessible through the internet. Much relevant data with respect to lots are also published on completion of the sale of auctioned property in *Estates Gazette*, which is a leading United Kingdom real estate weekly journal published in London.

Sometimes, some of the data from some of the variables are also sent in electronic format by the auctioneers to investors and property professionals listed in their databases.

With the possible exception of the data with respect to yields, all the empirical data is directly in the public domain. In addition, the largest of the London property auction houses also publicly lists the yields of lots sold. As far as the remaining lots are concerned, investors and property professionals calculate the yields from the published data on rents reserved and hammer prices. This is normal market practice to facilitate a transparent and efficient market. It therefore follows that all the empirical data relied upon in the research is in the public domain.

1.7 Summary

Sale-and-leaseback is an increasingly popular model in the United Kingdom for the splitting of operational property from core business activity. The principal British retail banks have identified the model as a preferred way of splitting their operational property from their main business of offering banking services to the public.

The main motivations behind the adoption of the sale-and-leaseback model can be found through the review of the literature. The literature review also highlights the trends in using the model for banking-halls and investors' reactions to the opportunities afforded by investing in retail bank premises.

Electronic searches indicate that no research has been undertaken into the sale-and-leaseback of British banking-halls. In this respect, there is no extant knowledge on British banking-halls as a medium of property investment. A gap, therefore, exists in the knowledge. In testing the hypotheses and answering the research question, this research fills that gap in knowledge. Filling the gap in the knowledge will equip investors and their professional advisors with the capacity to better maximize returns from British banking-halls as property investments.

A rich seam of data exists in the form of the published catalogues and results of banking-halls sold at auction. This seam has a large number of data sets and several variables that are capable of analysis.

Theory on the sale-and-leaseback model and the sale-and-leaseback of British banking-halls is inducted through a review of the literature. The research question and the hypotheses are formulated from the theory. Deductive analyses of the empirical data are then used to test the hypotheses. Through this process, answers are given to the research question, changes to practice are suggested and recommendations are made.

1.8 Structure of the thesis

Murray (2006: 123-124) and Trafford and Leshem (2008: 54-57) show the importance of giving an appropriate structure to a thesis. At the most basic level, chapter headings do give some indication of the general structure. However, Murray argues that those chapters, too, should be given their own structure. As far as the main structure of the thesis is concerned, this is based upon the individual chapters and the order in which they are written. Such ordering of the chapters provides an orderly development of thinking for both the researcher and the readers (Trafford and Lesham).

In order to given an overall view, to show the development of thoughts and to demonstrate the links between chapters, the structure to the current thesis is summarized in Table 1.2.

Table 1.2 Structure of thesis

Chapter Number	Chapter Heading	Purpose
1	Introduction	Introducing the topic and the research objective
2	Sale-and-leaseback: the common approach to investor participation	Examining the sale-and-leaseback model
3	Influencing factors of banking-hall investment: practical perspectives	Examining the provision of retail bank premises
4	Influencing factors of banking-hall investment: theoretical perspectives	Developing theory
5	The Research Design and Methodology	Identifying the way in which the research is conducted
6	The qualitative study: identifying the predictor variables	Identification of predictor variables
7	Introduction to the Statistical Analysis	Summarizing the data
8	Exploratory Data Analysis	Identifying the distribution of the data and likely relationships
9	Advanced Analyses I: Original dataset incorporating calendar data	Model-building based upon calendar date
10	Advanced Analyses II: Original dataset incorporating index data	Model-building based upon IPD UK retail index
11	Further validation from post-study data	Using those data from after the study period to further validate findings
12	Conclusions and Recommendations	Reflection on practice and research together with conclusions

2 Sale-and-leaseback: the common approach to investor participation

2.1 The sale-and-leaseback model and its evolution

Adams and Clarke (1996) describe the typical sale-and-leaseback transaction as being one in which the vendor simultaneously sells and property and takes a lease on that property from the purchaser. Usually, such transactions involve the vendor selling the freehold, or its legal equivalent in jurisdictions not subject to English law, and the simultaneous taking of a lease immediately subject to that freehold. However, it is possible for a sale-and-leaseback transaction to involve the selling of a superior lease and the simultaneous taking of a subservient lease by the vendor.

Sale-and-leaseback is not an entirely new model in the United Kingdom. It is a model that has been used in the past. The model had been used by a number of retail companies back in the late 1920s and early 1930s as a means of financing the expansion of the core businesses (Adams and Clarke, 1996). The sale-and-leaseback model had existed prior to 1987. However, the decision of the now defunct Woolworths Group plc to commence a process of sale-and-leaseback of its British operational properties in 1987 precipitated a trend in the United Kingdom. Since then a growing number of British listed companies have adopted the model, which has continued to remain popular (Northedge, 2005). The model has been adopted to a lesser or a greater extent by all the leading British retail banking companies. In recent years, the model has become widespread in many sectors throughout the United Kingdom (Northedge, 2005), including public houses, convenience stores, pharmacies, off-licences and petrol filling stations amongst others.

The hiving-off of operational real estate from the core trading activities of a business is now commonplace in the United Kingdom. Trading companies have been driven towards splitting their operational property assets from their core business activities. This means that businesses move real estate assets out of the ownership of the business. Sale-and-leaseback is not the only means by which such a split can be achieved. Alternative models that have been explored include creating separate property and trading companies within a group, sale-and-manageback, and outsourcing. Beyond the commercial sector, some public bodies have adopted the private finance initiative model. Investors and business operators alike need to evaluate which is the best method for the holding of any given type of operational property. However, where the principal five British retail banking groups have hived-off their operational properties, sale-and-leaseback has been the favoured approach.

. Notwithstanding the effects of accounting reforms under International Accounting Standard (IAS) 17 (McClary, 2010; McMillan (2011), the hiving-off of property has meant that it has been moved from the balance sheet and placed in the profit and loss account (Kingsmill, 2005). Hence, occupational leases replace ownership. Such a split may facilitate either the outright spinning-off of the property resource or the retention of it within a separate component of the same group (Northedge, 2005). Operational property is the property that a business needs to occupy in the conduct of its trade. The property is not part of the business activity. On the other hand, business activities are those activities normally undertaken by an enterprise with the specific purpose of making a profit.

The split of operational property assets from core business activities is driven by a small number of factors. These factors include finance, accounting, taxation, expertise and flexibility. They are not necessarily mutually exclusive. Due to the differing and changing priorities and motives of trading companies and their professional advisors, different strategies have been devised. These strategies continue to be subject to change as theory evolves and the regulatory framework adapts (Tipping and Bullard, 2007).

Long-established joint-stock companies usually held their operational property in the same companies as their core businesses during the century following the establishment of the first such companies. Often, the operational properties held by these companies had long since been paid for. Therefore, little regard was paid to the value locked into such properties except at times when these companies were placed under financial pressure. As a result, the equity locked into these properties was treated as free inputs into the businesses (Northedge, 2005).

Northedge states that the emphasis changed from the 1970s onwards when it became apparent that trading companies should separate their operational properties from their core business activities. Since then, a number of means by which operational property may be split from core businesses have been employed. In most cases, these have involved the placing of real estate assets in a separate company.

Sale-and-leaseback investment opportunities are continuing to present themselves in the United Kingdom. Not only are new sale-and-leaseback arrangements being created, but volumes of existing sale-and-leaseback stock are being traded as the original investors choose to liquidate their investments. This is equally true of British banking-halls as well as many other classes of operational real estate.

2.2 Structures for holding operational property

Sale-and-leaseback is not the only model for the holding of operational property. A business may hold its operational property and core business activities together as a single entity. If a division between the two is made, it can be done in a number of ways. Also, any such division can vary in the extent to which it separates the two. The extent of separateness is determined primarily by the structure that is adopted for the holding of operational property. However, the way in which any given structure is applied may also have an influence.

The principal structures through which operating companies in the United Kingdom hold their operational property are:

- A single company for core business activities and the holding of operational property.
- Division into separate property and operating companies owned within a single group company.
- Sale-and-leaseback.
- Sale-and-manage-back.
- Outsourcing.

These structures are not entirely exhaustive and are capable of adaptation to suit individual needs or a specific property asset class. Also, some premises have never been owned by the operating companies, but have always been rented from third party landlords.

2.2.1 Holding operational property and trading within a single entity

Until recent decades, apart from those premises that were always rented, the norm was for operational property to be held within the same legal entity from which core business activities were also conducted. Within the private sector, this usually entailed the two being held either by a sole principal, a business partnership or an incorporated company. From the end of the nineteenth century, it became the norm for larger, British businesses to convert to incorporated companies. This change did not become prevalent in smaller, family businesses until several decades latter. Presently, it is the norm for all but the very smallest of enterprises to be structured as incorporated companies with limited liability.

Where operational properties were held as an integral part of the trading business, in times of adversity they were treated as being part of the business's family silver that could be used to release capital. Often, the historic cost of the properties was a fraction of the existing value. Furthermore, over time many real estate assets had become debt free. In such cases, the directors of trading companies and their advisors frequently fail to appreciate and realize the present value of their properties.

Most trading activities can be expected to deliver better returns than real estate on capital (Whiting, 1999). Generally, companies can be expected to achieve a return of twenty percent per annum on capital deployed in normal trading activities. In contrast, yields on operational property cannot normally be expected to equal that level of return. Thus if the capital locked into operational property were to be re-deployed into core-business activities, companies could expect to become more profitable. Holding the two elements within a single company will distort a trading company's accounts. Some of the alternative structures can also be used to release capital from operational property for more efficient use in trading activities.

The placing of real estate assets in a separate company has been an end in itself. The managers and proprietors of many businesses have been content with such an arrangement, whereby the property company charges the operating company for the use of group real estate assets. This model has at least allowed the real estate assets to be treated as being capable in their own right of producing income from invested capital. It has also facilitated the better management of property assets through the employment of real estate professionals.

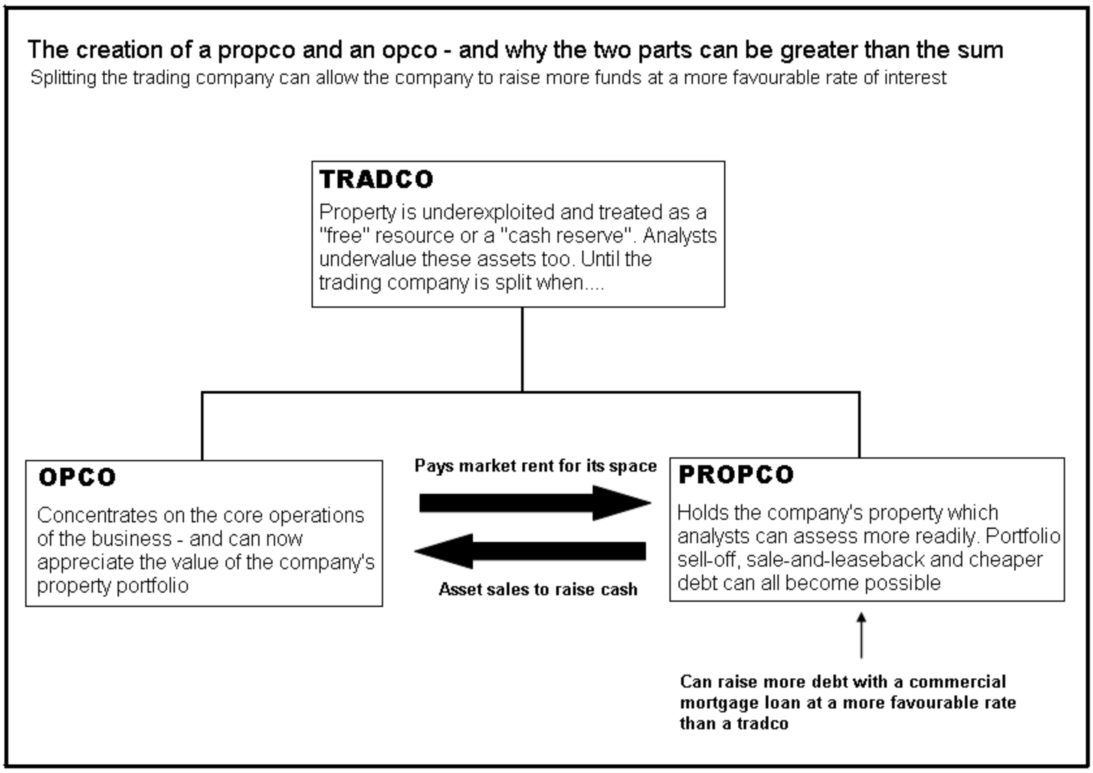
There are alternatives to creating separate trading and property companies under the umbrella of a single group. The hiving off of real estate assets into a separate property company can also be a necessary initial step towards the creation of more sophisticated arrangements (Tipping and Bullard, 2007). It is certainly the approach adopted as an initial step in the creation of sale-and-leaseback structures in the United Kingdom. For example, British retail banks that have entered into sale-and-leaseback arrangements have initially created group property companies into which they have placed operational property. The group property companies have then sold their legal interests in the properties to third party investors, thereby completing the sale-and-leaseback transactions.

2.2.2 The split between operational and property companies

Usually, the first step towards separating operational property from trading activities is to place each into separate companies within a company group. Greater focus on the component parts can be achieved through the division of the single entity into separate operating and property companies. These divisions are sometimes referred to respectively as *Opco* and *Propco*. Thus the provision of operational property by the property company to the operating property is made through the creation of an internal market. Through this structure, the operational company is expected to pay the property company the market rent for the property that it occupies. Such structure allows the operating company to focus its resources and expertise on its core business activities. Similarly, the property company is utilized to fully exploit the group's real estate assets.

As illustrated in Northedge's diagram in Figure 2.1, a simple division into separate operating and property companies that are wholly owned subsidiaries of a group is a first step (Northedge, (2005). The model has now evolved so that the separate operating and property companies need no longer be wholly owned subsidiaries within the envelope of a single group. One alternative has been for trading companies to enter into joint venture arrangements with third parties for the holding of their operational properties.

Joint ventures for the holding of operational property are normally established through structures that are sometimes known as Special Purpose Vehicles (SPVs). SPVs are normally in the form of companies that are jointly owned by the trading company and a specialist property company. Although normally based upon a company structure, SPVs are in essence partnerships between the two parties. However, the division of ownership is not necessarily divided equally between the two. The extent to which each party owns a share of the SPV is something that is agreed between them at the outset. Nevertheless, SPVs are a means by which operating companies may continue to retain some of the potential rewards and benefits of property ownership whilst simultaneously reducing the risk and liabilities.



Source: Richard Northedge

Figure: 2.1 The division of a trading company into separate property and operating companies

Tesco Plc has led the way with property joint ventures. In 1996 Tesco and British Land created BLT Properties as a joint venture (Northedge, 2005). In March, 2004 Tesco placed thirty-three stores and two distribution centres into a joint venture with Topland. During 2005 Tesco entered into two further joint ventures into which it placed some operational property (KPMG, 2006). The two structures entered into during 2005 had some of the attributes of the sale-and-leaseback model. Certainly, Tesco did sell legal interests in the properties concerned and did then take leases on those properties. However, these were not pure sale-and-leaseback transactions in the sense that the vendor simultaneously disposed of its legal interests and took back leases on the properties concerned. Instead, Tesco retained a fifty per cent stake in the SPV. Also, in one of the transactions, Tesco reserved the option to re-purchase the properties either after ten years or at the end of the lease. Rather than use a company as the SPV in the second transaction, Tesco opted to place the properties into Jersey Property Unit Trusts. Tesco has since opted to revert to using more conventional sale-and-leaseback structures (Chesters, 2006).

Securitisation has been an approach that the market has applied at times to those operators with high credit ratings. Since the onset of the credit crunch these have been much rarer, but Tesco was able to arrange three such deals during 2009 and the first half of 2010 (Thame, 2010). These were based upon Tesco keeping the freeholds, but granting long leases of between 99 and 999 years to the investor, taking back thirty year leases and then using the security to offer a commercially mortgaged-backed securitisation (CBMS) to the market (Thame). Following the onset of the credit crunch, only a company of Tesco's size and exceptional good credit rating would be able to induce investors to enter into such an arrangement (Thame).

The Opco/Propco split is capable of further refinement. It can be developed into a pure sale-and-leaseback transaction. Other models can also be developed. Other models that have recently been used in the United Kingdom are normally known respectively as *outsourcing* and *the Private Finance Initiative* (PFI).

2.2.3 Sale-and-leaseback

Sale-and-leaseback is the sale of a superior interest in a property on the basis that the investing purchaser lets it back to the vendor for an agreed term. In recent years, the sale-and-leaseback model has become increasingly used in the United Kingdom. Although the current trend towards sale-and-leaseback can be traced back to the model being adopted by Woolworths in 1987, it had previously been used by some United Kingdom retail companies in the late 1920s and early 1930s. Furthermore, the model had been used by small concerns and weaker covenants in the decade or so immediately prior to 1987. For example, a number of farmers adopted the model to raise capital from institutional investors during the 1970s and 1980s.

The widespread use of the sale-and-leaseback model has its roots in the United States. The first sale-and-leaseback deal in the United States was undertaken by Safeway Stores in 1936 (Adams and Clarke, 1996). This transaction consisted of the sale-and-leaseback of several supermarkets. This set the precedent for the widespread sale-and-leaseback of offices, factories and retail premises in the United States. The popularity of the model in the United States may have influenced its adoption in some European countries. For example, in the Netherlands it has been adopted by Vendex, in Sweden by the Coop, in Germany by both Dresdner Bank and Deutsche Telekom, and in Switzerland by the Union Bank of Switzerland (Tipping and Bullard, 2007). The Spanish bank Banco Bilbao Vizcaya Argentaria (BBVA) sold 947 retail bank branch premises in 2009 and a further 153 such premises in 2010 on sale-and-leaseback terms. While companies in the United States led the way in adopting the model, the motives of those in diverse European countries that followed suited may not have necessarily been the same.

The Woolworths sale-and-leaseback transactions took place between 1987 and 1991, during which time the model was applied to one hundred and twenty-nine of its properties. Boots the Chemist followed suit by selling fifty of its freeholds between 1988 and 1991. More recently in 2005, the company sold a further 312 of its stores in a single sale-and-leaseback deal with Reit Asset Management (Chesters, 2005a). Under the terms of this deal, Reit was prohibited from breaking-up the portfolio and gave Boots the option to vacate a limited number of premises without penalty. The purpose of the deal was to use the proceeds for the repayment of short-term loans and to go towards a three year programme of modernization and expansion of its premises. Since its adoption by Barclays Bank in 1992, the sale-and-leaseback model has been widely used by the British retail banks.

During the beginning of the twenty-first century, the sale-and-leaseback model has become much more widespread in the United Kingdom. The types of property to which the model has been applied in the United Kingdom include:

- Banking-halls
- Leisure properties
- Retail units
- Offices
- Manufacturing premises
- Government buildings

The list is far from exhaustive and all the types of property listed, other than banking-halls, are capable of sub-division, giving a wide range categories to which the model has been applied.

Some of the categories of sale-and-leaseback properties that have come to the market have done so through being sold as individual lots at public auction. Some categories have never been exposed to the market in this way. Instead, these categories have been sold as portfolios and by private treaty. Both approaches have been adopted for a number of categories.

There are some variations to the model. Recently, the most common variation in the United Kingdom has been for sale-and-leaseback transactions to include a tenant's option to break the lease early at certain, prescribed dates. Adams and Clarke give the example of the options of vendors and purchasers to require that properties be bought back from the purchasers at pre-determined prices as examples of variations of the model.

Irrespective of which approach is used to bring sale-and-leaseback properties to the market, two aspects have always been important in ensuring the success of these disposals. These two aspects are strong covenants and adequate lengths of lease. These two factors are necessary to facilitate mortgage funds. This is more especially so where individual properties are sold at public auction. Frequently, leases for a minimum of fifteen years are sought in such cases.

Prior to the start of the twenty-first century, most sale-and-leaseback deals were offered to the market on the basis of standard, institutional leases being granted. Such has been the popularity of the model amongst investors that vendors have been able to vary the terms to their advantage. Competition amongst investors has meant that vendors of sale-and-leaseback properties have been able to impose more restrictive terms. Vendors have been able to do so on the back of an extremely buoyant market that has seen yields falls as property investors have competed to acquire the freeholds.

The approach taken towards the sale-and-leaseback of the public houses of the recently defunct London & Edinburgh Swallow Group (LESG) is an example of high investor demand being exploited to the advantage of operating companies. Between early 2004 and its placing into administration in September, 2006 and subsequent winding-up, LESG sold several hundred public houses on sale-and-leaseback terms. Due to keen investor demand for its properties, LESG was able to avoid rent reviews to full market rental value for the first three reviews. Instead, the first three reviews, with the third occurring in the fifteenth year, were set to be based upon the compounding of a prescribed annual percentage rate five yearly. Originally, the LESG lots were offered on new thirty-five year leases with the prescribed rate of compounding set at 3½% per annum. This was very soon reduced to 3%. By the middle of 2005, it had been further reduced to 2½%. Similarly, during the same period the lease lengths of these lots were reduced from thirty-five to twenty years. The vendors and their advisors appreciated that buoyant investor demand would enable them to achieve higher hammer prices whilst simultaneously reducing future rent liabilities.

Other operating companies were similarly able to ensure that the terms of their sale-and-leaseback deals reduced their risk. For example, Your-Move.co.uk. Limited's sale-and-leaseback agreements provide for a tenant's break clause after ten years. Likewise, many of the banking-halls that Barclays Bank have sold on sale-and-leaseback since October, 2005 have provided for similar break clauses. Where the recently sold Barclays Bank retail premises have been sold on sale-and-leaseback with break clauses, these have usually been timed to occur after fifteen years (www.property-auctions.com). The adoption of break clauses by some operating companies is a clear recognition that the demand from individual investors is strong enough for those investors to accept less favourable terms.

Most of the sale-and-leaseback transactions in the United Kingdom have not restricted landlords' scope to a greater extent than the standard, institutional leases. However, some more restrictive patterns have started to emerge in a few cases. The sale-and-leaseback by Boots, the pharmacy chain, of 312 of its smaller stores sold in August, 2005 is an illustration. When these properties were placed on the market in April, 2005, Boots stipulated that any purchaser would be prevented from securitizing or dividing the portfolio (McClary, 2005a). Boots completed a sale-and-leaseback deal on the properties with Reit Asset Management four months later on that basis. The deal also gave Boots some flexibility insofar as it was permitted to vacate up to three per cent of the portfolio by value in any given year (Chesters, 2005a).

Not all sale-and-leaseback transactions have been in the form of individual properties sold at public auction. Often entire portfolios of real estate have been sold as a single entity by private treaty to investors. This approach has been popular in mainland Europe. In Germany Dresdner Bank sold €2 billion worth of its operational properties in on a sale-and-leaseback basis to Fortress in December, 2005 (KPMG). In the same month, Prakiker sold properties in Germany, Hungary and Portugal to IXIS AEW on the same basis for €500 million (KPMG). In the United Kingdom there have been some high profile disposals of portfolios of operational property on sale-and-leaseback bases. In 2004, British Land purchased sixty-five public houses from Spirit Group in a sale-and-leaseback deal (Cardew, 2004). Boots adopted such an approach with the deal with Reit Asset Management in August, 2005 in spite of having previously sold many other of its retail premises on sale-and-leaseback individually at public auction. Several of the principal British banks have sold some of their operational premises on this basis whilst simultaneously selling other units individually on sale-and-leaseback bases at public auction.

Although the sale-and-leaseback model has seen extraordinary growth in the United Kingdom since 1987, there have been several sale-and-leaseback failures during late 2005 and 2006. In October, 2005 S-Mart, the convenience store chain, collapsed, dragging down several private investors with it (Cruise, 2005a). Unwins, the off-licence chain, collapsed during the 2005 Christmas holiday (Cruise, 2006a). Helical Bar had acquired a third of Unwins' 380 retail premises in March, 2005 and promptly set about selling on the investments individually. During 2006, both the LESG and Provence public house groups collapsed. It remains to be seen if these failures will have any impact on investors' appetites for certain types of sale-and-leaseback transaction.

In a few instances, it is the operating companies that have become dissatisfied with the sale-and-leaseback model. Blackstone, a predator seeking to acquire the Center Parcs holiday village operator, announced that if successful in acquiring the business it would seek to reverse the existing sale-and-leaseback arrangement. Blackstone stated that the Center Parcs business had been burdened by the expensive arrangement (Gibson, R, 2006).

2.2.4 Sale-and-manageback

Another model for the splitting of operational property from core business activity is sale-and-manageback. This model has been increasingly applied to some types of operational property. It has mostly been applied to the leisure industry. In particular, it has been applied to a number of hotel properties. Even established United Kingdom hotel chains are using the sale-and-manageback model to separate their operational properties from their core business activities.

Sale-and-manageback can be defined as an arrangement whereby the investor takes control of the business as well as ownership of the property (Morrison, 2005). The operator is then granted a small proportion of the turnover and operating profits (Schäfer-Surén, 2005).

Presently, sale-and-manageback has been seen by the operators of businesses within the leisure sector as an alternative to sale-and-leaseback. This is because the sale-and-leaseback model is now perceived to have lost some of its attractions. The recent adoption of international accounting standards in the United Kingdom now means that future rental payments covenanted within leases have to be recorded as liabilities within balance sheets. Also, Stamp Duty Land Tax (SDLT) requirements now mean that longer leases in particular create greater tax impositions for business operators.

Schäfer-Surén (2005) states that billions of pounds of hotels have been sold by the large hotel chains in sale-and-manageback transactions in recent years. The thinking behind these arrangements is that the hoteliers consider that they can achieve greater returns on their capital if it is invested in their businesses rather than tied up in real estate. The sale-and-manageback model is a departure from sale-and-leaseback. Leases now have to be shown as liabilities on balance sheets. Sale-and-manageback does not create such a liability. On the other hand, Schäfer-Surén argues that management contracts may not deliver such good returns to hoteliers. He argues that it may take five or six hotels under management to have the same effect on profit and loss as two or three leased hotels or one owned hotel.

Management contracts may be a cost-effective way for new, under-capitalized hotel groups to enter the market. However, for established groups originally owning their own hotels, the model may be more of a short-term means of raising share prices.

Romney (2005) and McKenna (2005) have claimed that the sale-and-leaseback model has lost some of its lustre due to the increased burden of SDLT on lessees' interests. However, Imber (2005) states that occupiers who take leases as part of a sale-and-leaseback transaction are not subject to SDLT for those leases. This is confirmed by H.M. Revenue & Customs (2004).

Care does need to be exercised in the drafting of sale-and-manageback contracts. This is especially so with the need to make a clear distinction between tenant and licensee status with respect to the business manager. Often it can be difficult to distinguish between the two, whereas the law requires more precision. The law considers the essence of an agreement rather than the nomenclature. Therefore, it is quite possible for a management contract to give the opposite effect to that intended. If wrongly drafted, it may inadvertently create a lease under the Landlord and Tenant Act 1954. In the case of Bon Apettito v. Michael Poon (2005) the court held that a restaurant was subject to a lease in spite of the contract document describing the arrangement as a management contract (McKenna, 2005 and Romney, 2005). The result could be catastrophic for either party. From the new property owner's perspective, an error would mean that the operator would acquire security of tenure rights under the Landlord and Tenant Act 1954. This would be in spite of the intention of his being a licensee with reduced security rights. From the operator's perspective, an error might render it liable for SDLT. Failure to pay SDLT in such circumstance would be a criminal offence. Furthermore, both the property owner and the operator might discover unintended and disadvantageous accounting and taxation implications arising from an error.

Sale-and-manage-back has increased in use in the United Kingdom. However, great care needs to be exercised in order to prevent an unplanned sale-and-leaseback. Moreover, the manage-back model tends to be limited by property type in its application. Generally, it is confined to the leisure sector and more especially hotels.

2.2.5 Outsourcing

Outsourcing by a trading company may take more than one form. Indeed, even outsourcing of real estate requirements can be far from being a precise model. However, property outsourcing is defined by Kingsmill (2005) in terms of being a generic model as:

> "Where the management of the properties is also undertaken by the investor".

The history of outsourcing in the United Kingdom is well documented. According to George and Pazzi-Axworthy (2002), it started with the Ministry of Defence's transfer of the armed services' married quarters to Annington Homes in 1996. This transfer had more of the attributes of a conventional sale-and-leaseback deal. However, George and Pazzi-Axworthy do acknowledge that the first true real estate outsourcing contract in Europe was probably the PRIME project entered into by the United Kingdom's Department of Social Security in 1998. The public sector continued to enter into outsourcing contracts simultaneously and in conjunction with the Public Finance Initiative (PFI).

Only a limited number of outsourcing partners have been providing outsourcing structures in the United Kingdom. The two principal providers have been Mapeley Limited ("Mapeley") and Land Securities Trillium Group Plc.

Land Securities Trillium claims to have originated the outsourcing concept in the late 1990s (Land Securities Trillium, 2004). In its promotional literature, Land Securities Trillium gives the following definition to the concept and its own role in the process:

> "The solution enables an organisation to transfer its short-term and long-term property needs to a single specialist provider. We can then take on the ownership, management and development of all or part of a client's estate, enabling property assets and liabilities to be converted into an integrated property service. This releases capital for the client, delivers operational savings, provides occupational flexibility, reduces risk and provides price predictability."

One of Land Securities Trillium's most high profile outsourcing contracts was the one entered into in 2001 for the holding, managing and development of the operational estate of the British Broadcasting Corporation ("BBC"). That was to have been a thirty year partnership to enhance the efficiency of the BBC estate. However, within five years, the BBC dissolved most of its outsourcing contract with Land Securities Trillium (Cahill and

Lazarus, 2005). There were criticisms that the contract had cost the BBC considerably more than had been initially anticipated (Cahill, 2004a).

The outsourcing of operation property requirements has not been exclusive to the public sector. Mapeley's most high profile property outsourcing contracts have included the STEPS (The Strategic Transfer of Estate to the Private Sector) project. This project was established for the holding and management of diverse properties on behalf of H.M. Revenue & Customs (Mapeley, 2004). However, property outsourcing has also been procured from Mapeley by the private sector. In 2000 Mapeley entered into a twenty year property outsourcing contract with the bank, Abbey National Plc, with respect to latter's estate of approximately 1,300 properties comprising banking-halls and other premises (Mapeley, 2004; Wilding, 2010).

In 2001 British Telecom transferred the bulk of its estate to Telereal in an outsourcing deal (Dover, 2005a). This was undertaken in order to reduce British Telecom's massive debt (Devaney and Lizieri, 2004). In Deutsche Telekom's sale-and-leaseback deal, also transacted in 2001, it was to fund network expansion. Another telecommunications company, Cable & Wireless, abandoned the outsourcing of its operational estate at a late juncture.

In November, 2004, Abbey National was acquired by the Spanish bank Santander. During 2010, Santander undertook a review of the way in which it held the former Abbey retail banking premises (Wilding, 2010). By the end of September, 2010, Santander was reported as being close to re-acquiring the former Abbey portfolio from Mapeley (Wilding). The bank cited both achieving better value and the desire to be seen to show local commitment as reasons for owning operational property and hence the reversal of the outsourcing of the former Abbey operational property portfolio (Wilding).

2.2.6 Public Finance Initiative

Since the mid-1990s, the Public Finance Initiative (PFI) has been a favoured model by government and public sector agencies in the United Kingdom for the procurement of new operational buildings. These have included courts, schools, hospitals, prisons and tax offices (George and Pazzi-Axworthy).

PFI is not a model for the splitting of operational properties from core businesses. Rather, it is a device by which government and the public sector are able to avoid making capital payments for new public sector capital projects. Instead, capital projects are funded, provided and maintained by the private sector, which then takes regular payments spread over the long-term (The Economist, 2003).

2.3 Rationale behind splitting property from the business

The hiving off of operational property assets by operating companies in the United Kingdom is presently very much in vogue (Northedge, 2005). Where companies have hived off their real estate from their core business activities, they have done so for a variety of reasons. This, together with the diverse perceptions of trading companies and their advisors, has resulted in a number of different real estate strategies being adopted. Moreover, these strategies can change over time as theory develops and regulatory framework alters.

The sale-and-leaseback model is currently very much in vogue in the United Kingdom as a means of splitting operational property from core business activity. Furthermore, property investors have acquired an appetite for these investment opportunities where long leases, secure income flow and strong tenant covenant can be demonstrated (Kingsmill, 2005).

The principal motives behind the adoption of the sale-and-leaseback model are not necessarily mutually exclusive. They are:

- Finance
- Accounting
- Taxation
- Expertise
- Flexibility and liquidity

The distinction between property and the business conducted within that property is a relatively new concept. In the United Kingdom, it is one that arose during the twentieth century. Following the creation of the first joint stock companies, the normative model was for those companies to own both the core businesses and the properties. The directors and shareholders of these early companies were content to conduct everything through the one company. Over time, the operational properties of these companies became debt free through the redemption of any existing mortgages. Moreover, inflation had also eroded any outstanding debt. Therefore, over time these companies tended to treat their operational properties on an historic cost basis. The properties became treated as free inputs and reserves from which capital could be released during difficult trading conditions (Northedge, 2005).

It was in the United States in the 1970s that it was appreciated that trading companies needed to separate their real estate assets from their core business activities. It was perceived as being advantageous from a taxation aspect. The rationale has changed, especially in the United Kingdom. Whereas Charles Clore, a British property entrepreneur, built up the Sears group on the back of the capital values of its freehold stores, the approach is now different. In contrast, operational properties are now regarded as a means of unlocking capital for core businesses (Northedge, 2005).

Most structures for separating property from core business activity involve the placing of the freeholds and superior leases into separate companies. Such an arrangement may be an end in itself. The objective may be merely to create separate property and trading companies under the umbrella of a single holding company. However, this arrangement is also often a necessary step towards the establishment of other, more sophisticated structures. In particular it is a necessary first step in the creating of sale-and-leaseback deals. For example, Lloyds Bank, as it then was, formed Lloyds Commercial Property Investments Limited to facilitate the sale-and-leaseback of its banking-halls. Similarly, London & Edinburgh Inns Limited placed the freeholds of hundreds of its public houses into Jedi Inns Limited to facilitate the sale-and-leaseback of those properties respectively to two of its subsidiaries: Newlord Limited and Winlease Limited.

The principal motives behind the splitting of property from core business activities as listed above are examined as follows:

2.3.1 Finance

A trading company can optimise its property resources by releasing capital from that property for use in the core business. In so doing, the company might be able to use the released capital to achieve a higher return. Released capital may also be used to reduce borrowings. Sometimes, it is only possible to achieve one of these objectives. On other occasions, it may be possible to achieve both.

The directors and shareholders of trading companies need to consider the operational property as a business component having value in its own right. It is capable of being exploited in its own right. The property should be evaluated on a basis beyond mere historic cost. It should not be regarded as a free asset. Even the creation of an internal market between an operating company and a property company within the same group concentrates minds on the value of the estate. This gives the directors of the property company the option of selling or letting a property outside the group if they consider that that would be the optimal approach.

Such separation has financial reporting and borrowing ramifications. Trading companies measure performance in earnings per share. Property companies use net asset value. In terms of borrowing, a property company should be able to obtain cheaper sources of finance against its properties than an operating company could for non-property related loans.

An operating company should be achieving a twenty per cent return on its capital. Property yields do not even come close to that. Therefore, there is a case for releasing capital from property assets for use in core business activities where a much greater return should be achieved.

The splitting of property and operational activities into separate companies need not be an end in itself. The division can be taken even further. So, for example, the transfer of freeholds or superior leaseholds to a third party investor can be used to reduce the debt of the vendor group as a whole. Typically, such transfer can be achieved through sale-and-leaseback or through property outsourcing.

Many listed companies seek to retain their real estate assets as part of a strategy to protect themselves from hostile takeovers McClary (2006). However, in the case of the very largest companies, such a strategy is less warranted. Therefore, these companies are less restrained by concerns over takeovers from disposing of their real estate assets.

In contrast to McClary's (2006) argument, there has been a recent trend in the United Kingdom for takeovers to be funded through subsequent sale-and-leaseback deals. For example, immediately after acquiring the Little Chef restaurant chain in 2005, restaurant entrepreneurs Lawrence Wosskow and Simon Heath sought to finance the takeover through an abortive sale-and-leaseback deal with a consortium led by Prestbury Holdings (Chesters, 2005c).

Kingsmill (2005) cites Debenhams and the Big Food Group as examples where sale-and-leaseback featured as part of the funding used by predators in takeovers. In the United Kingdom, the Takeover Code provides that those making offers for companies must have the monies in place when so doing. Therefore, the norm is for lead banks to provide initial finance in the knowledge that they will be re-paid from the proceeds of subsequent sale-and-leaseback transactions (Kingsmill).

In the case of Wyevale Garden Centres, owning a portfolio of freehold operational property actually attracted hostile predators. The company had been targeted by property investors, who had identified opportunities. Some had identified re-development opportunities. Others thought that the company should capitalize on sale-and-leaseback opportunities (Dover, 2005d). Eventually, the company's board decided that it should dispose of some of its smaller sites, but that it would not enter into sale-and-leaseback deals for its retained operational property. The reason given for not proceeding with a sale-and-leaseback programme was 'its impact on operational gearing' (Dover, 2005e).

On the other hand, retaining some capital in operational property might well be less risky than using it all in trading activities. Devaney and Lizieri (2004) argue that equity markets might take the view that the holding of capital in real estate is less risky than employing it in its operational business activity.

On the downside the vendor loses an asset through sale-and-leaseback. The vendor is unable to release more capital from the asset either through another sale-and-leaseback or by borrowing against the property. Some see sale-and-leaseback deals as "selling the family silver". However, it is possible to devise means for vendors to continue to share in future rewards. Special Purpose Vehicles (SPVs) are a means by which a trading company can hive-off operational property whilst still benefiting from some of the future profits of the property. Tesco has been very pro-active in this area. It is also possible for sale-and-leaseback transactions to provide for options whereby a vendor may elect to re-purchase property at a pre-determined price (Adams and Clarke) or one determined on a prescribed basis. Such a course of action has to be weighed against the recognition that part of investor demand is buoyed by the anticipation of the investor benefiting from future growth.

2.3.2 Accounting

In recent years increased globalization has increased pressure for the standardization of how companies present their accounts in different jurisdictions. Previously there had been a distinction in the United Kingdom between what Lawson (2001) describes as operational leases for property and finance leases for other assets. This distinction was criticised by accountants on the grounds that operational leases were barely covered in accounts. Leased real estate was treated as being off-balance sheet (Royal Institution of Chartered Surveyors, 2011). This was in stark contrast to finance leases. For example, leased cars and photocopiers would be highlighted in company accounts as assets and liabilities, whereas a rented property would not. The G4 + 1 group, representing the accountancy professions of the United Kingdom, Canada, Australia, New Zealand and the United States, published proposals in 1999 for reform. Those seeking accounting reform argued that long-term lessees bore most of the risk associated in property ownership and that their leases should therefore be recorded in the accounts as finance leases in order to reflect this risk. Dover (2005b,) defines a finance lease as one that transfers the risks and rewards of an asset's ownership from the lessor to the lessee even if no legal transfer of ownership is made. She defines an operating lease as one where the risks and rewards of ownership remain with the lessor.

Reform did arrive in the form of International Accounting Standard (IAS) 17, which provides for the treatment of leased real estate mainly with respect to tenants and occupiers. IAS 17 requires that long property leases be shown as finance leases. It requires these to be capitalized and brought onto balance sheets as liabilities. In 2001 the European Union decided that International Financial Reporting Standards (IFRS), or International Accounting Standards (IAS) as they are now known, should apply from January, 2005 to all companies listed on any European stock exchange (Preston, 2004). Therefore, the European Union directive made the adoption of IAS 17 mandatory for all United Kingdom listed companies. This requirement is due to be extended to all companies. Proposed, reformed IAS 17 rules have been under discussion for a decade (Shah, 2010). The International Accounting Standards Board (IASB) was seeking to introduce another standard within a year or so of 2005 whereby all leased real estate would be treated on the basis of finance leases (Dover, 2005b). However, such implementation has been subject to delay. Not until August, 2010 was a draft agreement made for the International Accounting Standards Board to move towards implementation (Beddy, 2010). The final version is expected to be published by the IASB in the middle of 2011 (McMillan, 2011).

If implemented, the new IAS 17 rules will require lessees to show on their balance sheets by as soon as 2013 any lease liabilities (McClary, 2010). McMillan (2011) states that this is likely to be delayed until 2014, but with reporting based upon the accounts from 2012 onwards. Under the proposals, lessees will have to show lease payments during the accounting period in the profit-and-loss account and any subsequent liabilities under the lease in the balance sheet (Shah, 2010; Royal Institution of Chartered Surveyors, 2011; McMillan, 2011).

Since under IAS 17 all liabilities over the term of a lease will need to be capitalized, longer leases will be reflected by greater liabilities being recorded in the balance sheet. Such liabilities will be recorded on the basis of current or net present value. This will impact adversely on companies' gearing and profitability. Therefore, it is likely that those companies subject to IAS 17 will be more likely to seek shorter leases in order to reduce the impact (McClary, 2010; Shah, 2010; Vernham, 2011). Vernham suggests that lease terms exceeding fifteen years would become rare. Previously, many tenants had sought longer leases in order to spread out fitting-out costs. Now such tenants may seek break clauses, as lease liabilities will only be shown up to the break clause dates in the accounts. Lawson (2004) argues that landlords may seek to drive up the rents to offset the increased risk arising from accounting reform. However, it remains to be seen how the market would react to such pressures.

In the light of the reforms to IAS 17 proposed in August, 2010, some occupiers and their professional advisors are already questioning the continued usefulness of the sale-and-leaseback model (McClary, 2010; Shah, 2010). McClary (2010) and Williams (2010) argue that, if implemented, the reform will increase pressure from tenants for shorter leases and the greater use of break clauses. McClary (2010) also argues that in some cases the proposed reform might bring into question the benefits of leasing as opposed to freehold ownership by occupiers. Those occupiers who will be most affected by the proposals are those with multiple branches such as retailers and including retail banks (Beddy, 2010; Shah, 2010).

The real problem facing many companies is that the new accounting standard may increase their indebtedness beyond gearing levels agreed with their lenders. The risk is that some debts might be called in. Alternatively, lenders might seek to impose new, less favourable terms. Many trading companies are now looking for suitable structures to meet the challenges imposed by new accounting standards. IAS 17 had reduced the appetite of some companies to embrace sale-and-leaseback. Property outsourcing quickly evolved following the emergence of IAS 17. The model has caught on since the turn of the century, as property outsourcing has been treated as off-balance sheet. Hence the model has left gearing ratios unaffected. However, that, too, will be subject to reforms classifying property outsourcing as being finance leases and bringing it back onto the balance sheet.

Pressures for leases to be longer in some instances remain. Tenants saddled with fitting-out costs need more time to spread the costs. Also, the property investment market requires sufficiently long leases. In particular, property investors and their financiers require leases long enough to support the repayment of purchase costs. One solution is for the adoption of break clauses. Under the accounting rules, the rent liability of leases need only be calculated to the point of the first break clause. This is an approach that has been adopted by Barclays Banks for many of its properties sold on sale-and-leaseback since October, 2005.

It remains to be seen how the market reacts to the on-going accounting reforms. Dilution of the balance sheet through sale-and-leaseback might also affect a company's credit rating adversely. The real concern is that the capitalization of lease liabilities will show some occupiers' gearing ratios to be higher than that agreed with their lenders (Lawson, 2004; Beddy, 2010; Williams, 2010). These accounting issues are ones that need to be addressed. Nonetheless, sale-and-leaseback remains a popular model with both occupiers and investors.

2.3.3 Taxation

Adams and Clarke (1996) examined the implications of tax on sale-and-leaseback transactions. They found that taxation had had a different influence in the United States on the sale-and-leaseback model than in the United Kingdom. They referred to a study in 1990 of the effects of the model on the United States stock market. The study was for the period 1975-1986. That study found that large variations of marginal rates of corporation tax had a positive influence on share values where a company with a lower marginal rate leased property from one with a higher marginal rate. United States tax reform radically altered the position so that both overall tax and the gap between marginal rates reduced. The gap between the marginal rates was considerably reduced. Adams and Clarke argue that if a similar study had been conducted for the period following the 1986 tax reforms, the findings would have been very different.

The taxation system in the United Kingdom has been different from that in the United States. Therefore, it has had a different influence on how companies approach sale-and-leaseback from a taxation perspective. The reason for this is that depreciation is not tax deductible on real estate. An exception to this rule is the 100% capital allowance, which can be claimed for properties in Enterprise Zones. Otherwise, capital allowances may normally only be applied to plant and machinery within buildings (Adams and Clarke).

In the United Kingdom, lessees can benefit from tax relief on the properties that they rent. Lessees that are companies can benefit from offset rental payments against Corporation Tax. In this respect, owner occupiers are at a disadvantage, because only part of their mortgage payments can attract such tax relief. The part paid out as mortgage interest can attract similar tax relief. However, the element that represents the repayment of capital cannot attract such tax relief. Therefore, renting property clearly appears very attractive for those businesses wishing to maximize annual tax relief from their operational properties.

A liability for Capital Gains Tax may arise for increases in the value of property at the time of disposal. Such liability is calculated, after the deduction of any allowances, on the basis of any increase in value after having allowed for permitted indexation. Potential Capital Gains Tax liability can act as a deterence against sale-and-leaseback. Certainly, in the tax year in which properties are subject to a sale-and-leaseback transaction, the vendors could incur substantial Capital Gains Tax liabilities. Often, these liabilities can be reduced through the timing of the transaction.

Stamp Duty Land Tax (SDLT) is a tax that is immediately triggered by a transfer of a legal interest in real estate. Intra-group transfers in real estate interests are currently exempt from SDLT. In contrast, it is payable by third party property investors acquiring interests through a sale-and-leaseback transaction. However, the vendors of the superior interest in commercial property do not become liable for SDLT with respect to leases that they have entered into as part of the sale-and-leaseback transaction (H.M. Revenue & Customs, 2004).

2.3.4 Focus and expertise

The splitting of operational property from a trading company certainly permits the separate real estate entity to focus upon its property assets. The separate entity can call upon specialist, property expertise that might not be available within a trading company (Devaney and Lizieri, 2004).

Even the splitting of property and core business activities within separate companies within a single group, permits the focus and expertise. The separate property company can focus on the real estate assets. It can more easily engage specialist property expertise to optimise the real estate resource.

Sale-and-leaseback achieves a similar focus. Here, too, the investor is able engage specialist property expertise. Indeed, specialist property investors may have access to greater degrees of expertise than a group company. Outsourcing goes a stage further. Not only does it past on ownership of the property to a third party, but also the management of that property.

The ability to focus upon real estate assets and to exploit expertise may be a motive for placing a business's property into a separate vehicle, but it is not a reason in itself for choosing one model of separation over another. Other considerations are likely to determine which model is selected.

2.3.5 Flexibility and liquidity

Operating companies owning their own operational property have less mobility and liquidity. If the property is mortgaged, those companies are tied by the burdens of the debt.

Leasing property can also be a tie in the medium term. However, this issue can be addressed to some extent by the use of lessees' break clauses. Especially in more recent years, a number of the sale-and-leaseback transactions have featured tenants' break clauses. For example, some of the Threshers off-licences and the Barclays Bank premises sold in recent years have been subject to tenants' break clauses. Several of the Threshers properties were subject to tenant's break clause at five years. Similarly, some of the Barclays Bank premises sold since October, 2005 have been subject to tenant's break clauses after either ten or fifteen years.

Property outsourcing has been cited as a model that can deliver greater flexibility than sale-and-leaseback. Boots' disposal of 312 of its stores to Reit Asset Management in 2005 was in the main a sale-and-leaseback deal rather than an outsourcing one. However, it did give Boots greater flexibility insofar as Boots was allowed to vacate a limited number of properties (Chesters, 2005a and 2005b). Abbey National, the British retail bank, entered into an outsourcing agreement for all its operational estate comprising 1,320 properties with Mapeley. It transferred the entire estate to Mapeley, which was to supply both property and facilities management. The deal allows Abbey National to add and remove a limited number of properties from the estate. It is very flexible, but it is considered to have been very expensive (Devaney and Lizieri, 2004).

2.4 Sale-and-leaseback failures

Most of the sale-and-leaseback transactions sold through the London property auction houses have comprised properties considered to have strong covenants. Indeed, many of these lots have comprised properties where the lessee was to have been a leading, blue-chip company. Typically, the properties were to be let to FTSE 100 companies, banks and leading high street retailers. The high reputation of such tenants helped to drive the demand from investors for such properties offered at auction. Portfolios of sale-and-leaseback properties occupied by the same companies and others of similar high regard have also been sold to institutions and large investment companies by private treaty. Together, these transactions have helped to increase the profile and high reputation of the sale-and-leaseback model in the United Kingdom.

The increased profile and high reputation of the sale-and-leaseback model has helped fuel a growing demand from investors for sale-and-leaseback properties. This has allowed companies with a much weaker covenant to dispose of their properties by sale-and-leaseback. Some of these properties have been sold in portfolios by private treaty. Large numbers have been sold in recent years through the London auction houses.

Since the end of 1985, a number of the weaker covenant companies that had shortly beforehand used the sale-and-leaseback model failed. These companies had disposed of many of hundreds of their operational properties by sale-and-leaseback. The consequence of these failures was that hundreds of the sale-and-leaseback properties were left vacant with unpaid rents. To compound the situation, many of the properties were left in a dilapidated condition. Furthermore, they also tended to be over-rented so that if re-let they would be unlikely to achieve the former rents. Some investors have also been dragged down by these failures.

The high profile sale-and-leaseback failures in the United Kingdom include:

- London & Edinburgh Swallow Group of public houses and hotels
- Provence Commercial Properties public house chain
- S-Mart convenience stores (Cruise, 2005a)
- Unwins off-licence stores (Cruise, 2006a)

- Southern Cross care homes (Upward-only rents cripple Southern Cross, 2011)

During the credit crunch, some further failures were highlighted. These included Woolworths, which had been one of the first United Kingdom retailers to have disposed of its operational property by sale-and-leaseback. Even some of the stronger high street retailers are reported to be facing difficulties at present due to lower spending, changing retail patterns and competition from the internet. Therefore, more sale-and-leaseback failures remain a possibility.

2.5 A single opportunity for operators

Sale-and-leaseback is a model that allows operators to capture the benefits of such transactions with respect to their operational estates. However, such disposals of freeholds by operators can only be undertaken once. Put in another context, the family silver can only be sold once.

The current economic downturn has some similarities to the recession of 1989, albeit the present economic downturn has the extra dimension of a shortfall in liquidity and a lack of capital in banks' balance sheets. Gilbertson (2007) states that whereas most of the large corporate businesses owned the freeholds to their premises in 1989, they now no longer do so. Most of those companies lease their property now. Not only are those companies now unable to raise capital through sale-and-leaseback, but as Gilbertson states, they now have less asset value on their balance sheets. This is now compounded by the requirement of those listed companies to show leasehold interests as liabilities on their balance sheets.

2.6 Conclusions

Businesses in the United Kingdom, and indeed elsewhere, have become increasingly aware of the benefits of hiving-off their operational property assets from their core business activities. The sale-and-leaseback model is not the only means to this end, but the benefits of its application have become much more appreciated. Therefore, its application has become much more widespread. The use of the model especially by leading British companies has become much greater since the turn of the century.

As a result of its being the preferred approach for splitting operational property from operational business activities in the United Kingdom, sale-and-leaseback is the model that offers the greatest opportunity to those seeking to invest in such property. It is a process that cannot be repeated in the sense that a business can only sell-off its operational property once. However, new investors can purchase existing sale-and-leaseback properties from the current landlords. Sale-and-leaseback is certainly the approach that has been used to hive-off the vast majority of British banking-halls from the operational business of retail banking. The notable exception has been the outsourcing of the majority of the Abbey National estate. As the norm for investing in British banking-halls, the sale-and-leaseback model is the one worthy of being the focus of investigation.

Sale-and-leaseback has, at different times, been driven by a number of factors, including finance, accounting, taxation, expertise and flexibility. All these factors need to be taken into consideration when searching for the optimal way of holding operational property. Accounting reforms, requiring lessees to show the commitments arising from their leases as liabilities on their balance sheets, do not appear to have diminished the desire of British businesses to enter into sale-and-leaseback deals. However, it is finance and the need to optimize the use of capital that appear to be the principal influences driving the model's adoption in the United Kingdom (Tipping and Bullard, 2007).

3 Influencing factors of banking-hall investment: practical perspectives

3.1 Introduction

Retail banking has been subject to numerous changes and innovations during the last two decades in the United Kingdom as well as elsewhere. The sector has seen consolidation through acquisitions and mergers. New technologies in the guise of electronic communications, as well as electronic data processing and storage, have presented new opportunities and new challenges. Banks have also been under pressure from shareholders to boost profitability through the cutting of operational costs. The resulting attempts to cut costs have in turn put pressure on staff numbers and operational property needs.

The changes and innovations in retail banking have resulted in frequent re-examination of the operational property needs of the retail banks. Such re-examination is not confined just to the United Kingdom. It has also been widespread in several other parts of the world that have been influenced by consolidation and new technologies. It was predicted at the onset of these changes and innovations that there would be large-scale closure of retail bank premises as the pattern of banking business adapted (The Economist, 1993). However, bankers have subsequently realized that there is a benefit in retaining branches (The Economist, 1992b). Whereas branch numbers saw some initial contraction, they subsequently saw some expansion.

This chapter aims to investigate the factors influencing banking-hall investment through examination of the changes in retail banking and the associated needs of retail banks.

3.2 Review of the literature on banking-hall premises

3.2.1 Valuing banking-halls

Due to their unique attributes, such as vaults, cashiers' counters and automated teller machines (ATMs), retail bank premises are often valued using the contractor's method of valuation. Schafer (1994) argues that such an approach is not always appropriate, since some banking premises are capable of conversion to other uses. He favours capitalization of income. This means in United Kingdom property valuation terms the adoption of the investment method of valuation. Moreover, he states that present-day developers of banking premises design such premises with a view to possible, subsequent conversion to other uses.

Additional research in the form of a published paper (Tipping, 2006) is used to define the appropriate valuation techniques to be applied within the analysis of sale prices and rents reserved with respect to the banking-halls sold at public auction. As a result, this research also highlights gaps in the knowledge with respect to the existing valuation practice. These further gaps in the knowledge are also filled. Consequently, this research makes a significant contribution to professional practice by making recommendations with respect to the valuation of property investments. In particular it investigates the impact of the time value of money (Tipping, 2006).

On the other hand, Schafer's investigation, based upon the early 1990s, concludes that United States banking executives were then less concerned with market rents and market capital values when factoring-in the occupational costs of new operational premises. He argues, on the contrary, that most financial institutions subject to his study were more interested in the relationship between the bank deposits made at any given branch and the cost of owning or renting those branch premises.

Schafer's study is based upon New York. The study's analysis omits obvious outliers in order to reduce the risk of the results being skewed. Also, the study is focused most on the Manhattan part of New York for a number of reasons: the small geographic size of the area; the large number of rented bank premises; the homogenous economic and social spread in the area; and the consistency of rents in the area. Through the use of linear regression analysis, Schafer demonstrates a high correlation between the level of deposits made at branches and the rents passing on those branches. He concludes that banks are viewed as secure tenants with low vacancy rates and that this reduced risk has the effect of reducing yields on the properties.

3.2.2 Rationalization or expansion of the branch network?

Greenland (1995) shows that all the major banks have been re-assessing their branch networks. This has resulted in both branch rationalization and changes in branch inter-relationships and hierarchies (Greenland). Burton (1990) states that during the 1980s, retail banking in the United Kingdom underwent its most radical changes since the growth of branch networks following the establishment of the joint stock clearing banks in the mid-nineteenth century. In contrast, the changes during the 1980s saw a contraction of the branch networks. The decline in the branch numbers of the then leading four United Kingdom banks between 1978 and 1987 is shown in Table 3.1. Nellis and Lockhart (1995) give more up-to-date figures for branch rationalization. They state that bank branch numbers were further reduced by twenty per cent between 1982 and 1993. However, their figures are calculated on a slightly different basis, since from 1988 they include the branches of Abbey National, which had previously been a building society. Burton states that between 1970 and 1987, building society branch numbers increased between three and four-fold. This is corroborated by the findings of Greenland. This may have something to do with building societies starting to offer current accounts and other financial services during this period.

Table 3.1: Main clearing bank branches 1978-87

	1978	1987	% change 1978-1987
National Westminster	3,262	3,101	- 5%
Barclays	3,054	2,767	- 9%
Midland	2,466	2,127	- 14%
Lloyds	2,325	2,162	- 7%
Total	**11,107**	**10,157**	

Source: Committee of London and Scottish Clearing Banks 1988 (Burton, 1990)

Since Burton researched her paper in 1990, retail bank branch networks have continued to be placed under pressure. Post office banking, supermarket banking and technological advances in electronic communications have all been cited. Furthermore, banks have been under pressure to reduce staff and property occupation costs.

Greenland draws on what is known as Christaller's central place theory. This theory seeks to explain not only the geographic spread of branches in the distribution of services, but also the relationship between different branches. Branches are placed in a hierarchical order. In adapting this approach, Greenland places banks' networks in a hierarchy ranging from stand-alone ATMs to the banks' national headquarters. Adoption of the theory allows banks to reduce costs through avoiding duplication and through referring customers to appropriate centres for the provision of specialist financial services. This has inevitably impacted upon the banks' operational property needs and made some locations redundant.

However, not all changes to the bank branch network have resulted in a reduction in numbers. Against the trend of branch closures, some United Kingdom building societies and smaller banks have embarked upon a programme of branch expansion (Greenland, 1995). He cites the examples of Abbey National and Yorkshire Bank having expanded their branch and ATM coverage in order to extend their presence over a greater geographical area.

The process of branch closures continued into the new millennium. For example, Barclays closed 171 British branches on the same day in 2000. This particularly high-profile act attracted a great deal of adverse publicity for Barclays (The Economist, 2001). Other banks also attracted criticism for their closures. 'The Economist' (2001) states that banks are now much more aware of the public-relations aspects of large-scale branch closures. Notwithstanding that there have been some closures, the pace of any consolidation has been much slower than that predicted and retail bank premises have not been made redundant by technological innovations (Duxbury, 2010).

Evidence from Australia suggests that some banks have more recently sought to reverse branch closures by opening new ones. The major Australian bank, Westpac, has re-opened branches as a means of improving business, especially in the provision of mortgages (Jones, 2003). Bank of Queensland also embarked on a branch expansion policy in recent years and cited this as being the factor behind improved profits (Fraser, 2003). In the United Kingdom, HSBC has decided to open some new branches (Chesters, 2006 and Poulter, 2006). However, these branches are to take on the attributes more akin to retailing than banking. During mid-2006, the bank announced that it was seeking modern, retail-orientated premises in prime retail locations (HSBC Holdings plc, 2006).

3.2.3 Post office banking

The United Kingdom's Post Office has remained a government-owned enterprise. Much of its monopoly has now been removed, but it still does operate a vast branch network, albeit in reduced numbers. The Post Office previously owned Girobank, through which it operated money transfer activities. However, a strategic decision was subsequently made to sell Girobank to Alliance & Leicester (The Economist 28 Oct 1995).

Banking and money transfer activities have often been undertaken by post offices throughout the world. Although the United Kingdom's Post Office had previously taken the decision to dispose of Girobank, the provision of banking and financial services has since seen a renaissance in the post office sector. This has become the case in several European countries. In the Republic of Ireland the Irish post office, An Post, offers banking services throughout its branch network on behalf of Allied Irish Banks (AIB). This is facilitated by customers' use of AIB Banklink cards and AIB credit cards at An Post branches (Allied Irish Banks, 2005). In Spain, the national post office, Correos, likewise offers banking services as agents to Deutsche Bank (Deutsche Bank, 2005a, 2005b and 2005c). A slightly different approach is taken in France. The French post office, La Poste, allows customers to open La Poste accounts, which can be accessed remotely (La Poste, 2007).

In a reversal of its previous decision to sell Girobank, the United Kingdom's Post Office has decided to follow the trend of other European post offices by re-entering the market for the provision of banking and financial services. This arose out of frustrations arising from the British government's failure to privatise the Post Office as a result of public and parliamentary opposition. Therefore, the British government gave Post Office Counters, a subsidiary of the Post Office, greater commercial freedom in the running of the 20,000 branches that it then operated (The Economist, 1995). Post Office Counters was already undertaking money transfer activities on behalf of Girobank, which it had previously sold to Alliance & Leicester. It therefore seemed logical for Post Office Counters to seek to undertake the receiving of deposits and the making of payments for other banks as well. Barclays, having initially rejected an approach, subsequently changed its mind when its competitor, the then Lloyds Bank, was approached by Post Office Counters (The Economist, 1995). There has since been a queue of banks seeking to participate. Post Office

Counters offer automated banking services for each of the principal five British retail banking groups as well as for several of the others (The Post Office, 2003).

Since opening up its branch network to United Kingdom and Irish banks on an agency basis, the Post Office has also decided to take a more direct role in the provision of banking and financial services. Therefore, in 2002 it entered into a joint venture with Bank of Ireland to create Post Office Financial Services (POFS) (The Banker, 2004). The Post Office and Bank of Ireland each took a fifty per cent stake in POFS. This arrangement enabled Post Office customers to seek financial services from Bank of Ireland subsidiary, Bristol & West, including the taking out of loans ranging from £1,000 to £25,000 (The Post Office, 2005). The Post Office also started to offer three and five-year equity bonds on behalf of the Bank of Ireland (The Times, 2005).

The Post Office has also sought to enhance its banking services through the provision of Automated Teller Machines (ATMs). In 2006 the Post Office agreed to the provision of 100 new, non-fee charging ATMs by Royal Bank of Scotland and National Westminster Bank in sub-post-offices in deprived areas (O'Grady, 2006). Also, the Post Office decided to update its postal orders in 2006 to make them more like a cheque and thus capable of being processed through bank clearing systems (Flanagan, 2006).

Although the Post Office has done much more to engage with banking and financial services in recent years, its branch network has been shrinking. The branch network had diminished from 21,000 in 1985 to 14,000 in 2006 (O'Grady, 2006). The government announced that branch numbers were to be reduced by another 2,500 by the year 2009 (*The Post Office network can never be absolutely cost efficient*, 2006; *Our post offices will still need protecting,* 2006).

Post office banking has grown in the United Kingdom in recent years. Withdrawals from selected accounts of most British banks can now be undertaken at Post Office branches. This may have been seen as a further threat to the most marginal of retail banking premises. However, the reduction of Post Office branches must now partly offset that threat.

3.2.4 Supermarket banking

Steinborn (1994) put a case for operating small banking units within supermarkets. She argued that such units could capture more business and that such in-store banks cost only a fraction of the price of conventional banks to build. She also pointed out that in-store units could break even on much smaller deposits. During 1994 there were 2,191 in store bank branches in the United States, which accounted for three per cent of all branches (Steinborn).

At the close of the twentieth century, Australia saw a rush to establish retail banking through supermarkets (O'Connell, 1998). This was pioneered by the Commonwealth Bank of Australia in a tie-up with Woolworths, which had 680 stores throughout Australia at the time. The strategy was to focus on branded transaction cards. In the United States, the supermarket in-store bank branch has been an approach that has often been adopted (Flynn, 1997). This latter approach was adopted in Australia when ANZ Bank and Coles Myer established some in-store banking kiosks (O'Connell).

According to O'Connell, the attractions of a supermarket tie-up for the banks were: potential sources of new customers; customers' trust in supermarket brand names; low operating costs; and the exploitation of loyalty schemes.

The in-store branch model has not been adopted in the United Kingdom, but British banks have used the supermarkets in another way to increase market exposure. Instead, they have teamed up with the larger supermarkets to create supermarket-branded banks. By 1998, this approach had been extensively embraced in the United Kingdom. Sainsbury's Bank plc was a joint venture between Bank of Scotland, which owned a forty-five per cent stake, and the Sainsbury's supermarket chain, which owned a fifty-five per cent stake. Having been established in 1997, Sainsbury's Bank achieved a customer base estimated to be about 700,000 within twelve months (O'Connell). Tesco Finance was established as a joint venture, with equal shareholdings, between Royal Bank of Scotland and Tesco plc, the largest supermarket chain in the United Kingdom (O'Connell). Tesco then proceeded to consolidate its position by buying-out Royal Bank of Scotland's interest in the venture in July, 2008 (White, 2008). By May, 2011, Tesco Bank had opened six branches and was set to open more (Kivlehan, 2011).

Both the Sainsbury's and Tesco experience show that sizeable numbers of people, perhaps at a time of customer dissatisfaction with the traditional retail banks, have been prepared to adopt the model of the bank and retailer joint venture where a strong retailer brand name exists. It is noticeable that the joint venture partners of both supermarket chains were at the time barely represented in the south-east of England. This suggests that the banks concerned may have been seeking to extend market exposure in those regions at the time. The supermarkets are motivated by other factors. The Tesco supermarket banking operation in the United Kingdom offered higher savings and lower loan rates. The additional costs associated with that needed to be absorbed by the supermarket, which could hope that the model of rewards in the form of discounted supermarket products, plus penalties for cash withdrawals, would make the model work for them (O'Connell).

3.2.5 New entrants to UK retail banking

During 2010, it became very apparent that new banks were pitching to enter the United Kingdom retail banking market (Kelly, 2010). Dissatisfaction amongst customers, especially in the wake of the credit crunch, meant that they were looking for new entrants to the market to offer longer opening hours and more consumer-friendly terms (Kelly). In March, 2010, a new British bank named Metro Bank was granted a banking licence (Wallop, 2010; Metro Bank, 2011). By May, 2010, Metro Bank had entered into fifteen year leases for two new branches (*Metro banks two more locations*, 2010). These premises were respectively at Fulham Broadway underground station in London and Borehamwood in Hertfordshire (*Metro banks two more locations*, 2010). Metro Bank's business plan was to open four branches by the end of 2010, forty branches by 2014 and two hundred branches by 2020 (Metro Bank, 2011; Kivlehan, 2011). The business model is to situate branches within the M25 close to London Underground stations and for these branches to be customer-friendly venues open for longer hours and for seven days a week (Metro Bank; Kivlehan). The concept that underpins Metro Bank's model was created by Vernon Hill (Metro Bank), who had previously successfully set up a similar new bank in the United States.

Metro Bank is not the only potential new entrant into British retail banking (Cave, 2010). The American bank, J.C. Flowers, has been negotiating a tie-up with the Kent Reliance

Building Society as a means of entering retail banking in the United Kingdom (Kelly, 2010). Sir Richard Branson, through his vehicle Virgin Money, unsuccessfully sought to purchase Northern Rock with a view to entering retail banking (Kelly). There was speculation that Virgin Money was still looking to enter retail banking in the United Kingdom (Cave, 2010; Aldrick and Ebrahimi, 2010). By 2011, Virgin money was pitching to buy six hundred Lloyds Banking Group branches that the European Commission had wanted removed from the group (Kivlehan, 2011).

3.2.6 Innovation in the branch network

Due to the high costs of staffing and accommodation, retail banks had hoped to persuade their customers to switch from conducting their business through branch premises to conducting it through electronic media. Advances in new technology facilitated the introduction of ATMs. This was followed by the introduction of EFTPoS (electronic funds transfer at the point of sale) technology (Prendergast and Marr, 1994). These technologies are now widespread. Greenland (1995) states that those ATMs placed within branches exist to enhance service and efficiencies rather than to replace the branches. However, he acknowledges that in places, these ATMs will also require rationalization, having met saturation point.

Telephone banking and, more recently, internet banking have become another means by which customers can manage their accounts. Branchless internet banks have also developed in the United Kingdom. However, although a significant number of customers have been willing to conduct some business on-line, they have insisted on retaining branches (The Economist, 2001). According to Dover (2006b), bank customers use telephone and internet banking in tandem with the traditional branches rather than instead of them.

3.2.7 The profile of the modern retail banking premises

Technological changes have reduced staff numbers and the space required in bank branches for normal account functions (Greenland). These changes have resulted in a transformation in processing and telephone enquiries from branches to central locations. Both Greenland

(1995) and Mathison (2001) show that bank branches have seen a shift from being locations for merely undertaking account transactions to places offering retail type services. As a result, bank branches now typically offer additional services such as mortgages, insurance and financial products designed for businesses. The banks have exploited the opportunity to cross-sell these and other products.

The shift in the functions of bank branches has inevitably influenced changes in the design, layout and size of branches. According to Greenland, the traditional branch design and layout is not always conducive to the present desire to achieve the necessary retail image. Capita Symonds (2011) state that there is a need for retail bank premises to become inviting, customer-friendly venues rather than traditional, solid buildings, which are coming to be seen as old-fashioned. No doubt this has something to do with HSBC's decision in mid-2006 to open retail-type branches for the purpose of creating suitable environments for cross-selling to the public (Chesters (2006), Poulter (2006) and HSBC Holdings plc (2006)). A similar approach had already been adopted by Abbey National (The Economist, 2001). It is also the approach to branch design adopted by the new entrant to the market, Metro Bank (Kivlehan, 2011). Greenland summarizes the key points affecting the design of modern bank branches. Those points impacting on the physical design of the premises are:

- The focus is on retailing and cross-selling rather than traditional banking activities conducted behind bandit screens
- Corporate branding in terms of décor and layout for each branch
- Glass frontages instead of ornate, stone facades
- Open-plan layouts
- ATMs in lobbies accessible twenty-four hours a day

Although the design, layout and size of branches have been under pressure to change, the high street remains the favoured location for British banking-halls. According to Dover (2006b), in 2006 HSBC had planned to spend £400m on extending its presence in the high street in the following years and HBOS had announced its intention of opening a further 100 branches. On the other hand, she states that Barclays would be consolidating its position on the high street by closing branches of its Woolwich brand.

3.3 The sale-and-leaseback of banking-halls

The present trend for the sale-and-leaseback of operational property from the estates of leading United Kingdom companies was started by Woolworths in 1987. The model was soon adopted by several of the leading British retail banks. Cavanagh (2003) provides the data for the sale-and-leaseback of British banking-halls at auction between 1992 and 2001. Barclays Bank was the first to sell some of its retail banking branches on an individual basis at public auction, when it sold 61 properties between 1992 and 1994. What was then Lloyds Bank followed suit in 1996. The newly merged Lloyds TSB Bank sold 445 branches on a sale-and-leaseback basis between 1996 and 2001. Midland Bank, before it became HSBC, sold 163 branches on a sale-and-leaseback basis between 1997 and 1998. Early British banking-hall sale-and-leaseback deals are summarized in Table 3.2.

Many of these early auctions focused on small lot sizes, which made them particularly attractive to small, private investors (Healey & Baker, 1997). Other factors that were highlighted as making the investments attractive were:

- The covenant strength of a leading British bank
- Attractive term of years on the leaseback
- Full repairing and insuring covenants
- Upward only rent reviews
- No Valued Added Tax on the purchase price

(Healey & Baker, 1997)

The smaller banking-hall lot sizes attracted many private individuals, who were new to both property investment and the property auction process. These new investors frequently outbid institutions and established professional investors to acquire these smaller lots. These lots were within the purchasing capabilities of the small investors. They also generated yields that exceeded the return on capital placed on deposit at the time (Cavanagh, 2003).

A rich seam of data exists for those British banking-halls sold individually at public auction on sale-and-leaseback terms. These data include variables that ought to be capable of testing in order to ascertain any influence that they may have on investment decisions. From the freely available data, it ought to be possible to test the relationships of:

- Yield
- Region
- Lot size
- Identity of tenant bank

(Source: www.propertyauctions.com)

The effects of branch cross-over may also have some influence upon investment decisions. However, data with respect to this variable are not available from the databases of published auction results. Furthermore, the limitations of resources with respect to this research prevent the measurement and recording of such cross-over.

Table 3.2: Early British banking-hall sale-and-leaseback deals

Bank	Year	Number of lots	Total Realization (£)	Average Yield (%)
Barclays Bank	1992-1994	61	25,600,000	8.00
Nationwide	1993	7	6,772,500	7.70
Lloyds TSB Bank	1996-2001	445	151,072,000	7.50
Midland Bank	1997-1998	163	22,032,000	7.80
Total		**676**	**205,476,500**	

Source: John Townsend, Cushman & Wakefield Healey & Baker (includes some deals involving the then Edward Erdman) (Cavanagh, 2003).

The process of selling retail bank branches on sale-and-leaseback has become popular throughout much of Europe. For example, the Swiss Bank, UBS, sold eighty-six of it branch premises on sale-and-leaseback terms in November, 1999 (Pottinger et al, 2002). Fortress purchased €2 billion worth of Dresdner Bank's operational premises in Germany on a sale-and-leaseback basis in December, 2005 (KPMG). Deutsche Bank sold much of its operational estate in Germany through sale-and-leaseback deals in November, 2003 and December 2004 (Deutsche Bank, 2004 and Euro Property, 2004). In October, 2006, Bank of Ireland agreed to sell a portfolio of thirty-six of its retail branches in the Republic of Ireland on sale-and-leaseback terms for €240 million (Hipwell, 2007). In April, 2007, Bank of Ireland announced that it would be seeking to sell its fourth tranche of retail branches, this time comprising twenty-nine of its retail branches (Hipwell, 2007). The Spanish bank Banco Bilbao Vizcaya Argentaria (BBVA) entered into sale-and-leaseback arrangements on 1,100 retail branches between 2009 and 2010.

The United Kingdom has a mature and established investment property auction market. This auction market is focused in London, where all the prime auctioned property is sold through a small number of specialist auction houses. Collecting data from the websites of the specialist auction houses shows that a very large number of British banking-halls have been sold on sale-and-leaseback by auction. In the rest of Europe the investment property market is not so well established. Therefore, it is more common to see such properties sold by private treaty. However, sometimes the leading banks have sold some of their operational property outside the auction room. This has usually been reserved for the largest lot sizes and has involved portfolios of premises being sold en bloc. The Royal Bank of Scotland has in the main avoided the auction process in the disposal of operational premises on sale-and-leaseback terms. Instead, it sold 300 branches to Ackerman Group for £100m in 2005, and during 2006 was seeking a purchaser for a second portfolio (Dover, 2006a). Barclays Bank trialled the sale-and-leaseback of a small number of retail bank premises by auction in late 2005. The bank sold a large number of branches on sale-and-leaseback at auction during 2006 and 2007 (Thomas, 2006a, Cruise, 2007 and Phillips, 2007). Furthermore, Barclays Bank also sold a portfolio of twenty-four banking-halls to Flodrive for £67m in September, 2006 (Cruise, 2006b). HSBC, when branded as Midland Bank, sold 163 banking-halls on sale-and-leaseback between 1997 and 1998 (Cavanagh, 2003). However, HSBC has recently sold branches by both private treaty and auction (Cruise, 2006b).

The disposal of entire portfolios of operational property on sale-and-leaseback terms is an approach much used in United Kingdom. Many of the principal retail banks have sold entire portfolios of their operational property en bloc to single investors in this way. Even those banks that have sold branches individually at public auction have adopted this alternative approach for other parts of their estates. Examples during 2006 alone include:

- Barclays Bank's sale-and-leaseback of 15 prime branches to Prudential for £85m (Thomas, 2006b)

- Halifax Bank of Scotland's £72m sale-and-leaseback of 15 branches to an Irish investor
- Barclays Bank's sale-and-leaseback of 24 branches to Flodrive for £67m (Cruise, 2006b)

Abbey National adopted a different model for separating its retail banking premises from its main business. Instead of adopting a pure sale-and-leaseback approach, Abbey National chose property outsourcing (George and Pazzi-Axworthy, 2002). The bank entered into its property outsourcing agreement with Mapeley. Under this arrangement, Abbey sold its entire estate to Mapeley from which it took leases of up to twenty years on the properties, but with a built-in flexibility to vary the lengths of the terms and to vacate at pre-determined prices (George and Pazzi-Axworthy). According to Mapeley (2004), the arrangement gave Abbey certainty and flexibility. Abbey's approach is not one that has been favoured by the other British retail banks. Outsourcing deals have since come under scrutiny on the basis of their expense and some have even been cancelled (Cahill, 2004a; Cahill and Lazarus, 2005). During 2010, Santander, as the new owner of Abbey, was seeking to reverse much of Abbey's outsourcing arrangement with Mapeley (Wilding, 2010).

International Accounting Standard (IAS 17) has had the effect of encouraging tenants to seek shorter lease lengths and of creating the greater use of break-clauses (Dover, 2005b). Leases entered into as part of a sale-and-leaseback transaction are not subject to Stamp Duty Land Tax (SDLT). However, when banking-halls are re-let on new leases, SDLT is likely to be a reason for banks seeking shorter terms.

3.4 Summary

In developing an understanding of the optimal ways of investing in British banking-halls, it is imperative that investors and their professional advisors acquire knowledge about those factors likely to influence the future operational property requirements of the leading banks. They need to know how the banks value their retail branches. They also need to understand how branch networks are likely to grow or contract; and they need to know the effects of the supermarkets, post office banking and innovations in the branch networks, whether in terms of technology or building type. Together, these factors should provide an understanding of the likely profile of the types of British banking-halls to be demanded by the leading banks in the future.

The factors that are expected to influence investment in British banking-halls are:

- Yield
- Region
- Lot size

These factors are based upon variables that produced empirical data capable of observation and recording. However, there remain other factors that may also influence investment. These factors include:

- The term of years of the leaseback
- Any discrimination between premises let to the leading banks and those let to the smaller banks and building societies
- Branch cross-over

Repairing covenants often influence property investors' decisions. All of the British banking-halls that have been sold on sale-and-leaseback terms have been subject to tenants' full repairing covenants. Although this factor has no doubt proved attractive to investors, a lack of alternatives to this norm means that it is not possible to use empirical data to test the effects of repairing covenants. Similarly, the norm of upward-only rent reviews every five years is an attraction to investors. Again, the absence of alternatives means that this factor cannot be subjected to scientific testing.

Strategic changes to the provision of retail banking services in the United Kingdom have seen radical changes to the branch networks with substantial branch closures, the spread of ATMs, changing branch hierarchies with new branch functions and changing branch appearances (Greenland, 1995). Investors and their professional advisors need to understand where these changes will have the most adverse effects. They need to know where these changes will either bring about branch closures or perhaps result in premises for which they would not be prepared to pay as much rent as they would for alternative premises.

Clearly, there have been pressures from the 1980s for branch rationalization and for changes between inter-branch relationships. As a result, many branches in marginal locations have been lost. Other locations may well become marginal. On the other hand, some of the leading retail banks are now looking to invest in new branches (Dover, 2006b and 2006c). Investors need to be aware of these trends, as they will influence banks' desire to occupy particular premises. Also, the pressures for shorter leases may have an impact. There will be higher demand for those banking-halls that best meet the profile presently sought by banks. This will have some influence on the optimal investment opportunities in British banking-halls.

4 Influencing factors of banking-hall investment: theoretical perspectives

4.1 Introduction

Kane and O'Reilly-De Brün (2001) argue that most research needs to be directed by theory. They liken theory to a map in terms of giving direction to deductive research (Kane and O'Reilly-De Brün, 2001: 38-39). However, research undertaken in the context of a professional doctorate is directed by practice, underpinned by the theoretical perspectives (Klenowski and Lunt, 2008). Hence, theory is capable of directing research. Such direction is important in order to ensure that the research is capable of answering the research question rather than being directed down blind alleys.

Collis and Hussey (2003: 122 and 357) define theory as being:

> "A set of interrelated constructs (variables), definitions and propositions that presents a systemic view of phenomena by specifying relationships among variables with the purpose of explaining natural phenomena."

Kane and O'Reilly-De Brün expand on Collis and Hussey's definition of theory by stating that theory is also capable of being tested.

In order to attain a scholarly dimension to the research rather than just to produce a descriptive study, it is necessary to engage extensively with the theories of other researchers through the extant literature (Trafford and Lesham, 2008: 72-75; Kane and O'Reilly-De Brün, 2001: 39).

One way to commence a theoretical study is to review the literature. Murray (2006) shows that a literature review can be adopted for a number of reasons. Each reason depends upon the nature of the study. So although a literature review is capable of being a stand-alone research project, it may form part of a larger study (Murray). However, a literature review is also a useful mechanism in informing and directing research. Huberman and Miles (2002) argue that a literature review is essential for developing theory from existing concepts, theories and hypotheses. They make the case for the inclusion of both conforming and conflicting literature. Both equally have a role in informing and directing the study. The inclusion of conforming literature with like findings is important, as it connects concepts, theories and hypotheses not necessarily previously connected with each other (Huberman and Miles). This can result in theory that has a:

> "Stronger internal validity, wider generalizability, and higher conceptual level" (Huberman and Miles, 2002: p: 25).

On the other hand, the inclusion of conflicting literature enables a greater understanding of the area of research. In adequately addressing conflicts highlighted in the literature review, the researcher can boost confidence that the findings of the study are correct (Huberman and Miles).

Trafford and Leshem (2008) go further by showing that theories generated from the literature continue to generate greater understanding through a process of summary, synthesis and analysis. Engagement with the literature facilitates the development of the theoretical perspectives that underpin the research (Trafford and Leshem).

Thought needs to be given to the literature, or *corpus*, to be relied upon. Trafford and Leshem show that the corpus is much more than the physical format in which it is presented. Rather, it is published knowledge that is capable of informing and directing the research (Badley, 2010; Trafford and Leshem, 2008). Trafford and Leshem rank types of literature within the corpus by dividing it into a hierarchy of three distinct groups. The highest in the hierarchy is that comprising original work that has made a major contribution to knowledge, whereas the others comprise little or no contribution (Trafford and Leshem). Trafford and Leshem also make the point that monthly professional journals often contain articles written by experts in a given field. They make a case for including such literature in the corpus, because they are often more up-to-date due to taking less time to reach publication than literature found in peer-reviewed academic publications.

4.2 Sources from within the corpus

An extensive search of the literature, dating from 1992 until 2009, was undertaken. This was done with a view to providing theoretical underpinning to the specific research question, which is:

“How can property investors select freehold banking-halls that are likely to maximize the return on their investment?”

Such return is defined as the initial yield. From the theoretical perspectives it should be possible to identify individual research questions with a view to selecting the most relevant factors in order to answer the specific research question. The search undertaken found no indication of any such literature on British banking-hall investments with respect to that period. It failed to locate findings of any previous research into factors showing an effect on the yields of British banking-halls sold by auction as investment properties. The retail sector is one of the three main types of commercial property. Banking-halls are in turn a type of retail property. Hence, banking-halls are a sub-type of commercial property. Notwithstanding the failure to locate findings from previous research relating to factors showing an effect on the yields of banking-halls, a search did locate a corpus of published works relating to factors influencing the yields of generic commercial investment properties situate in the United Kingdom. The sources of the literature ranged from peer-reviewed academic journals to reports produced by respected professional practices, and from property investment-related reference books to articles in professional journals. The articles from the professional journals are divided between those written by professionals expert in the field and those written by journalists.

Disregarding articles written by journalists, the remaining sources from the corpus are capable, on the basis of Trafford and Leshem’s hierarchy, of being separated into primary, secondary and tertiary sources.

4.3 Early models

Property markets are imperfect. In his study into the efficiency of the property market, Evans (1995) establishes that it is a heterogeneous market. Early theories defining rental values of agricultural land were set out respectively by Ricardo (1815) and Von Thünen (1826). Ricardo makes the assumption that rental value is determined by the fertility of the land in question. Von Thünen develops the concept of the isolated state, whereby the crop and the rental of outlying agricultural land are determined by the city at its centre.

According to Evans (1995), the implication arising from Ricardo and Von Thünen together is that the value is a function of fertility and location. However, such modelling is too simplistic to fit reality. Such modelling fails to take account of complex transport infrastructure, geography, the land having previously been used for other crops or uses and the city and its environs not being entirely self-sufficient (Crosier, 2007). Notwithstanding the limitations of the Ricardo and Von Thünen models, location, whether at parochial or national level, must have an effect on a property's rental and capital values (Tipping and Lam*,* 2010).

However, Ricardo's theory of fertility and Von Thünen's theory of location remain capable of application in a broader, more abstract sense (Evans, 1995). Fertility is capable of being broadened to include rent, capital value and yields for properties other than agricultural land. As Evans (1995) shows, modelling for properties other than agricultural land becomes much more complex due to the presence of buildings, which in turn comprise additional variables. Furthermore, each category of property type may have different forms of variables as having the greatest influence (Evans, 1995). Hence, for example, an office building's variables might include the availability of lifts and proximity to car parking, whereas those appertaining to petrol-filling stations might include proximity to the main highways and areas of widespread car ownership.

4.4 Identifying from the historic literature factors affecting yields

The corpus does include findings derived from research into commercial investment property in the generic sense. Ambrose and Nourse (1993) conclude that capitalization rates appertaining to property ought to be a function of property attributes and returns from other investment media. Ambrose and Nourse (1993: 221) defines *capitalization rate* as:

"The ratio of stabilized annual net operating income to purchase price."

Hence, capitalization rate is different from yield inasmuch as the former is calculated after management costs and allowance for voids. Notwithstanding that *yield* does not take into consideration management costs and voids, the findings of the Ambrose and Nourse study can be used in developing theory about influences on *yield*. Ambrose and Nourse merely adopt a stock market approach rather than a property investment approach in calculating relative returns on the respective investments. This difference in approach to calculation does not negate the findings of how factors influence the dependent variable, whether it is *capitalization rate* or *yield*. Hence, it is perfectly justifiable to include the findings with respect to the Ambrose and Nourse study, and any similar study using capitalization rates, in the corpus. Their study makes a good starting point in developing theory, because they claim that unlike previous studies, theirs does discriminate between property types in modelling returns.

The property attributes identified by Ambrose and Nourse as influencing returns are:

- Property type in terms of use
- Region
- Location, where location is defined by the narrower, more parochial definition.
- Size as represented by the amount advanced in the form of a mortgage to facilitate the acquisition of the property.
- Time

(Tipping and Lam, 2010)

Having identified both *location* and *region* as main factors influencing returns, Ambrose and Nourse find that the effects of *location* are insignificant, but add the proviso that it is possible that such a finding may arise out of the way in which *location* was recorded. Hence, there may be issues of measure validity with respect to this variable. With respect to *time*, Ambrose and Nourse find the different property types within their study perform differently over time.

The findings of Ambrose and Nourse suggest the factors that may be further considered and discussed with a view to developing theory. It is established by Ambrose and Nourse's study that property should not be treated generically when analyzing returns and hence yields. Property is not a single class, but an asset type comprising many categories. It is now established that real estate comprises distinct types which should be analyzed individually. Ambrose and Nourse (1993), and others such as Chen, Hudson-Wilson and Nordby (2004) and MacGregor and Schwann (2003), show how property type should be considered a factor influencing *yield* in the United Kingdom (Tipping and Lam, 2010). These researchers rely upon the first level of this factor which is the use of the property. In defining use, they lump all retail properties together as one type. No distinction is made between, for example, between pharmacies and banking-halls. However, each of these uses falls within a separate sub-set of uses within the overall retail property type. Therefore, it is logical that each sub-set, including banking-halls, should also be treated as separate factors. The matter becomes more complicated when considering individual occupiers. If it follows that pharmacies and banking-halls should be treated as separate factors, then surely there is a case to be made for treating different occupiers as different categories. In which case, the particular banking company occupying a given property should be considered as having a possible influence on the dependent variable.

Ambrose and Nourse do discuss both location and region. Although they recognize the uniqueness of a specific location, which Evans (1995) conceptualizes by location having several dimensions (Tipping and Lam, 2010), Ambrose & Nourse group each separate location into an appropriate region in their study. Ambrose and Nourse's study into capitalization rates examines nine types of commercial property investments in the United States of America. Using data extracted from the *Investment Bulletin* published by the American Council of Life Insurance, Ambrose and Nourse grouped each location into one of five regions for the purpose of analysis. The current research into the investment yields of British banking-halls sold at auction is similarly able to use data published both individually by the respective auction-houses and collected by the researcher using the www.propertyauctions.com website. The available data with respect to the British banking-halls show both the individual locations and those regions grouped into ten geographic regions.

Property professionals in the field of property investments sold at auction in the United Kingdom often refer to *lot size* and its significance. Such reference to lot size is not in relation to the physical size of the premises, but rather in relation to its sale price or value. In building their model, Ambrose and Nourse include *amount* as the sum outstanding on the mortgage advance of a property to fund its purchase. This is because their study is into the capitalization rates of different types of commercial investment properties over different regions in the United States of America arising from mortgage contracts. The current research is not into returns achievable form mortgage contracts. Instead, the current research is purely an investigation into yields achieved by one type of property, namely British banking-halls, over the study period with a view to predicting future yields of that property type. Therefore, if some form of monetary sum is to be included as a predictor variable in the current study, it needs to be an appropriate measure. Clearly, the nature of the current study dictates that the measure of the monetary sum adopted by Ambrose and Nourse in their study would not be appropriate. It is more appropriate to adopt measures used in professional practice for categorizing *amount* when dealing with the sale of investment property in the United Kingdom. Baum et al (2006: p. 76) suggest that the price could have an effect on yield. They certainly cite liquidity and the availability of monies as a possible influence. In which case, a distinction needs to be made between lower and higher-priced properties.

In conducting their study, Ambrose and Nourse apply different tests over time. Contrary to expectations, the first test in their study did not produce the anticipated results. Dismissing the results arising from this test on the basis of validity, they apply an alternative test based upon a cross-sectional time-series approach, which generates results that confirm previous analyses and theory. The important point here is not the outcome of the Ambrose and Nourse study, but rather to recognize the inclusion of time as a factor when building and analyzing models with respect to yields on investment property.

Ambrose and Nourse's study was of factors influencing capitalization rates for a number of categories of commercial investment properties in the United States of America. They claim that their study was the first of its type to categorize the types of property. These categories did not include banking-halls. The study was published in 1993 and was based upon data appertaining to transactions between 1966 and 1988. Some studies had previously been conducted using the same data source, but with the purpose of addressing other research questions. Previous studies into capitalization rates using the same data are criticized for their shortcomings by Ambrose and Nourse. For example, they highlight that a previous study by Nourse had omitted to account for any effect that may have been caused by property type. Ambrose and Nourse find, in their critique of the previous studies, that there is a suggestion that property returns are influenced by property type. Thus, in their own study, they categorized properties by user type. That being the case, where data exist for banking-halls, these premises should in any study be categorized as a separate user type. Logic would suggest that such types must be capable of categorization into further sub-types. It therefore follows that banking-halls might be further categorized into sub-types as defined by the brand operating at a specific branch.

4.5 Current theories of factors affecting yields

Earlier studies, such as that of Ambrose and Nourse (1993), suggest what the main factors affecting yield might be. Even further back, the works of both Ricardo (1815) and Von Thünen (1826) make an important contribution to thinking. Although both developed models relating to the rent of agricultural land and hence its capital value, their works are important inasmuch as they establish the importance of location.

More recent literature generates theory based upon more current thinking with respect to the factors affecting property investment yields. More specifically, there is more recent literature that relates to the performance of property investment yields in the United Kingdom (Dunse *et al.,* 2007; Baum, Mackmin and Nunnington, 2006). The early literature underpins the more recent literature, which in turn is capable of being used to derive theory.

4.6 The impact of investment diversification

Investment incurs risk. Such risk occurs both in property investment and other investment media. Fraser (2004) distinguishes investment risk between *specific risk* and *market risk*. Thus, whereas *market risk* is that risk incurred by market fluctuation (Fraser, 2004), *specific risk* is that risk incurred by factors appertaining to a particular investment or a particular class of investment. This distinction applies no less to property than to other forms of investment. Enever and Isaac (2002) argue that risk can be reduced through diversification. They state that an investment portfolio is truly diversified when at least ninety-five per cent of its fluctuation is attributable to the market.

Specific risk can be diminished and in some cases eliminated through diversification (Fraser). Good investment practice, therefore, dictates that an investment portfolio should be sufficiently diversified and balanced in order to reduce risk and its adverse consequences.

Louargand (1992) shows how Modern Portfolio Theory (MPT) developed techniques in investment diversification in order to reduce the effects of risk. Extant literature on MPT suggests not only how there might be diversification across different asset classes, but also within real estate as an asset class through investment in different property types (Louargand, 1992). However, Isaac and Steley (2000: pp 116-117) contend that most United Kingdom investors are unable to avoid specific risk with respect to property. This is because, due to the relative high capital cost of acquiring investment property, most investors are unable to purchase sufficient properties with which to achieve adequate diversification.

By its very nature, diversification necessitates taking selected variables with a view to ensuring that investment is spread either across different categories or across a range when dealing with continuous data. Hence, it follows that the way in which MPT is applied to diversification is likely to have some influence upon the factors that might influence yield.

4.7 The application of Modern Portfolio Theory

Fraser (2004: p 115) shows how investment risk can be reduced, although not completely avoided, through diversification. He cites the example of stock market investment where investors invest in portfolios of shares in different companies rather than a tranche of shares in just one company. He argues that by spreading their investments throughout a multitude of companies, stock market investors are able to reduce their exposure to risk. However, risk cannot be entirely eliminated. Fraser explains this by making the distinction between *Specific risk* and *market risk*. Specific risk is that risk which appertains to a particular investment. An investor is, as Fraser shows, able to eliminate specific risk by spreading his or her exposure across a wide range of investments. Market risk, also known as *systemic risk*, is another matter. Returning to Fraser's stock market example, he argues that this type of risk cannot be eliminated, because it arises from fluctuations in the market as a whole. Over time, overall market fluctuations will arise as a result of the macro-economic cycle.

Isaac and Steley (2000) show that the expected return from an investment is to a very large extent a product of the level of risk involved. This applies whichever measure of return is adopted, including initial yield, which is the normative measure in the United Kingdom.

Back in 1952, Markowitz produced an investment model whereby specific risk seemed to be capable of reduction by the selection of investments with a diverse range of attributes within a portfolio (Isaac and Steley). Markowitz (1952) shows that through diversification, the variance between the expected performance and the observed performance within an investment portfolio can be expected to be diminished. He argues that it therefore follows that a diversified portfolio can be expected to be superior to a non-diversified one. The prudent investor should seek to reduce adverse risk whilst at the same time investing in a portfolio that gives the highest return. Markowitz describes a portfolio that simultaneously achieves both of these aims as being an *efficient portfolio*. Hence, specific risk can be reduced by diversification within a given portfolio (Isaac and Steley). Such diversification within an investment portfolio is known as *Modern Portfolio Theory* (Fraser, 2004: p 115). Having originally been utilized in the stock market, Modern Portfolio Theory is now applied in property investment insofar as investors are capable of diversifying beyond the individual property (Isaac and Steley).

Markowitz argues that diversified portfolio investment through informed selection, rather than through mere speculation, is the best means of achieving an *efficient portfolio*. Statistical analyses are capable of devising overall strategy in the building and management of the portfolio, but informed judgment then needs to be applied to refine the portfolio in order to provide the best results (Markowitz). The reason for this is that statistics alone are not capable of taking account of an individual factor that might influence yield or any other measure of return. This is especially so with real estate, where each property is unique.

It follows that good practice should dictate that any investment portfolio, including a property investment portfolio, should be suitably balanced in order to reduce adverse risk. Institutional investors seek to address this by investing in as many asset classes as possible. As part of this process, they often have prescribed maximum and minimum levels of exposure in any one class of investment. Thus, a prescribed proportion of investment by institutions may be in property whilst the remainder will be placed in other asset classes. Specialist property investors, by their very nature depart from the norm of also investing in other asset classes. However, they are still capable of applying the principles of Modern Portfolio Theory by diversification through investing in categories of property that have a wide range of factors.

The normative model was that investors would critically analyze the market prior to establishing a strategy for investing their capital with the objective being capital maximization. This normative model assumes that investors will act rationally. However, this normative model has been challenged on the basis that it was based upon observations of large, institutional investors to the exclusion of small property companies (Gallimore, Hansz and Gray, 2000). Gallimore *et al.* (2000) show, in contrast to the normative model, that property investment decisions are not always taken rationally:

> "The general dominance of rationalist approaches has been challenged by behavioural decision theory, drawing from cognitive psychology, and addressed at closer examination of process features, including the decision environment and individual differences between decision makers (Tversky and Kahneman, 1974)".

The model relied upon by the conventional perspective assumes that investors select their investments on the basis of anticipated future returns on the invested capital. In the context of real estate, these anticipated returns may be calculated on the basis of the rental income, rental growth and capital growth either singularly or in combination. This normative model assumes that investors act rationally to maximize the returns available for the perceived level of risk attached to the investment. Most of the extant literature relies upon the behaviour of the large, institutional investors, which rely upon the advanced analytical procedures adopted by their research departments, underpinned by statistical methods, to determine their investment strategies (Gallimore *et al.*, 2000).

Notwithstanding that the normative model assumes rational behaviour by investors to maximize returns, the reality with respect to some investors' behaviour can be different from that suggested by theory (Gallimore *et al.,* 2000). Gallimore *et al.,* (2000) show that irrational behavioural interpretations can result in investors departing from the normative model that the large, institutional investors generally adopt. Hence, according to Gallimore *et al.,* (2000) a model of investment selection based upon irrational behaviour, based upon reality instead of theory, can be more appropriate since it seeks:

> "to represent what actually happens when a decision is made rather than prescribing a theorised model of decision making."

Thus, Gallimore *et al.,* show that behavioural theory can explain why investment decisions may be skewed away from those that might be expected if the normative model were adhered to. As a result, specific risk may be increased due to investment decisions having been taken on the basis of personal prejudice rather than on a rational basis. In the context of real estate, such personal prejudice may well result in undue bias in favour of a particular property or property class. As with previous economic downturns, the recession that followed the 2007 credit crunch has especially highlighted the pitfalls of irrational investment decisions. Recessions compound the effects of such behaviour by adding market risk to specific risk. From Markowitz (1952), it is clear that an *efficient portfolio* is best achieved by applying informed judgment supported by statistical analyses.

Modern Portfolio Theory (MPT) assumes a move from the general to the specific. By deductive reasoning, this means that initially the classes and weighting of each class are determined. The precise selection of investments within each category will only be undertaken once the general strategy has been determined (Lee and Byrne, 1998). Those factors that should form part of a MPT investment strategy are shown by Lee and Byrne as follows:

> "Typically investors in real estate have attempted to diversify portfolios through a process of naïve diversification, although MPT has been advocated as a more rational approach to the construction of real estate portfolios (Lee, 1992). One approach which arises is whether investors should remain in one region and seek diversification by real estate type within the region, or diversify across regions but remain within the real estate type? Another related issue is whether diversification by real estate type or region alone produces significantly worse results than full diversification by both real estate type and region?"

Hence, there are different approaches that may be adopted in seeking the diversification of property portfolios. However, the two respective studies of institutional investors' property diversification strategies in the United States of America, undertaken by Webb (1984) and Louargand (1992), find that the normative methods of property investment diversification are by property type and region (Lee and Byrne, 1998). Modern Portfolio Theory is applied for the diversification of property investment in the United Kingdom. However, as Isaac and Steley, 2000: 116-117) show, due to the high cost of property investment, most property investors in the United Kingdom are unable to diversify sufficiently to eliminate specific risk.

Although Modern Portfolio Theory is capable of being applied to property, it must be done appreciating the effects of covariance. Although Markowitz makes the case for portfolio diversification, he shows that not all risk can be eliminated through diversification.

This is especially true for market risk. However, Markowitz cautions beyond this toward specific risk inasmuch as the effects influencing yields may be correlated. Using his example of the stock market, Markowitz warns that not all variance can be avoided, because many of the factors influencing the performance of equities are inter-correlated. This is also true of real estate. Since the present research is into factors having the greatest effect on the yields of banking-hall investments, there is going to be some level of inter-correlation, especially within such a narrowly defined asset class.

4.8 Risks affecting property investment

Adopting a top-down approach, this literature review of the factors affecting property yields starts with early, simple concepts, and drawing on other fields develops more complex and up-to-date concepts. Although Ricardo's theory of fertility and Von Thünen's theory of location are too simplistic, Evans (1995) demonstrates their usefulness in the abstract sense. As a result, it is possible to deduce that factors appertaining to the attributes of a property will affect its yield. Other fields suggest what those attributes might be. Similarly, studies focused on the stock market suggest how *efficient portfolios* might be assembled. Academic sources, such as Dunse *et al.* (2007), and Baum, Mackmin and Nunnington (2006) relate the corpus specifically to property and indicate which factors are likely to have the most influence on property yields. However, transactions in the real world are conducted by investors and their professional advisors. Therefore, it is necessary not only to review literature from academic sources, but also to draw on the appropriate professional sources.

Even though, as Gallimore *et al.* (2000) show, property investors can and do act irrationally, French (2004) argues that the Investment Method of valuation has become a much more scientific process. The Investment Method of valuation is the one that is normally applied to investment properties (Enever and Isaac, 2002: p 65). French argues that it has become a much more rational method of valuation whereby value is determined by using comparables of similar properties in the locality sold around the same time as the valuation date (Tipping and Lam., 2010). If French's assertion remains correct, then reliance must be placed upon those variables that affect the attributes of individual properties the most in terms of investment (Tipping and Lam, 2010). It follows from what French says that location and time are factors that need to be considered.

Although French (2004) does point to locality being a significant factor in influencing property yields, he states that this needs to be in combination of properties having similar attributes. However, properties may have similar attributes in a number of different ways and at different levels. Therefore, it is important to identify those similarities that are likely to produce the most significant effects. Much of the more recent corpus focuses most heavily on property type by region. Although their study was based upon data relating to the United States of America, Ambrose and Nourse (1993) led the way in showing that the property type and region combined are significant in influencing yield. The two factors were shown to work in combination. Other studies, such as those by Guy and Henneberry (2000) and Hoesli, Lizieri and MacGregor (1997), which support such a hypothesis, have followed.

Hoesli *et al.* (1997), for example, undertook a study of factors of influencing property yields in the United Kingdom with a view to determining the best diversification strategies. Their study was an investigation of region by three generic property types: retail; office and industrial. They used two analytical techniques. They used cluster analysis on the basis that such analysis had been used in previous real estate studies to support diversification by region, but added the proviso that the technique can be too deterministic. The other method was to categorize cases by region. This was done on the grounds that previous studies in the United States of America had found distinct regions there for property performance. Therefore Hoesli *et al.* undertook analyses on the basis of eleven regions in Great Britain. Northern Ireland was omitted. The eleven regions adopted by them loosely correspond to the ten regions adopted by the leading London property auction-houses. The main difference being that Hoesli *et al.* divided what those auction-houses continue to define as *North-east* into *North* and *Yorkshire and Humberside*. Hoesli *et al.* not only treated Scotland and Wales as individual regions, but also sub-divided England on the basis of the nine Government Office Regions (GOR) adopted by Central Government for the sub-division of England, which were to form the basis of regional assemblies until these were shelved in some parts. These Government Office Regions form the basis of economic planning. The map of the GOR classifications for the period subject to the study and produced by the Office for National Statistics (ONS) is reproduced in Figure 4.1. The GOR regions also form the basis of the regions adopted by the main property auction-houses, except that *North-east* and *Yorkshire and Humberside* are both grouped together under the label *North-east*. Hoesli *et al.* argue that a greater part of yield, as opposed to rental value, will be common both across regions and property type in the United Kingdom. This is because macro-economic factors such as Bank of England interest rates, have a major influence. The study by Hoesli *et al.* did not find a clear regional effect on property investment yields. Rather, their study did find a difference in yields within a given region. Notwithstanding their not finding significant differences in yield between regions, as defined by their eleven categories of region, they did find a significant variation between what they described as *super-regions*, or what for clarity might be labelled as *provinces*. They found that these super-regions, or provinces, fell into three categories: London; the south of England, including East Anglia; and the rest of the country. The Hoesli *et al.* study found that yields were influenced by property types across three super regions based upon London and a north-south divide. The suggestion that such a north-south divide might exist is supported by Guy and Henneberry (2000), who argue that institutional investors deliberately under-invest beyond southern England.

Notwithstanding the arguments put forward by Guy and Henneberry (2000), Lee and Byrne (1998) and Hoesli *et al.* (1997) supporting a north-south divide within the United Kingdom by the three main property types with respect to commercial property investment yields, Enever and Isaac (2002) did not necessarily support such a hypothesis. Enever and Isaac argue that the retail sector does not show a great north-south divide. Though they do acknowledge that notwithstanding the absence of a significant north-south divide in terms of growth and rent, there can be variations between locations within any given region. However, caution needs to be applied when considering the arguments of Enever and Isaac, because the absence of a north-south divide in terms of growth and rent does not necessarily preclude a north-south divide on the basis of yield.

In 2007, the University of Sheffield's Department of Geography produced a map of the perceived north-south divide in Great Britain (Dorling, 2011). It placed in to the south those areas roughly corresponding with East Anglia, London, the South-east and the South-west regions as defined by the Government Office Regions.

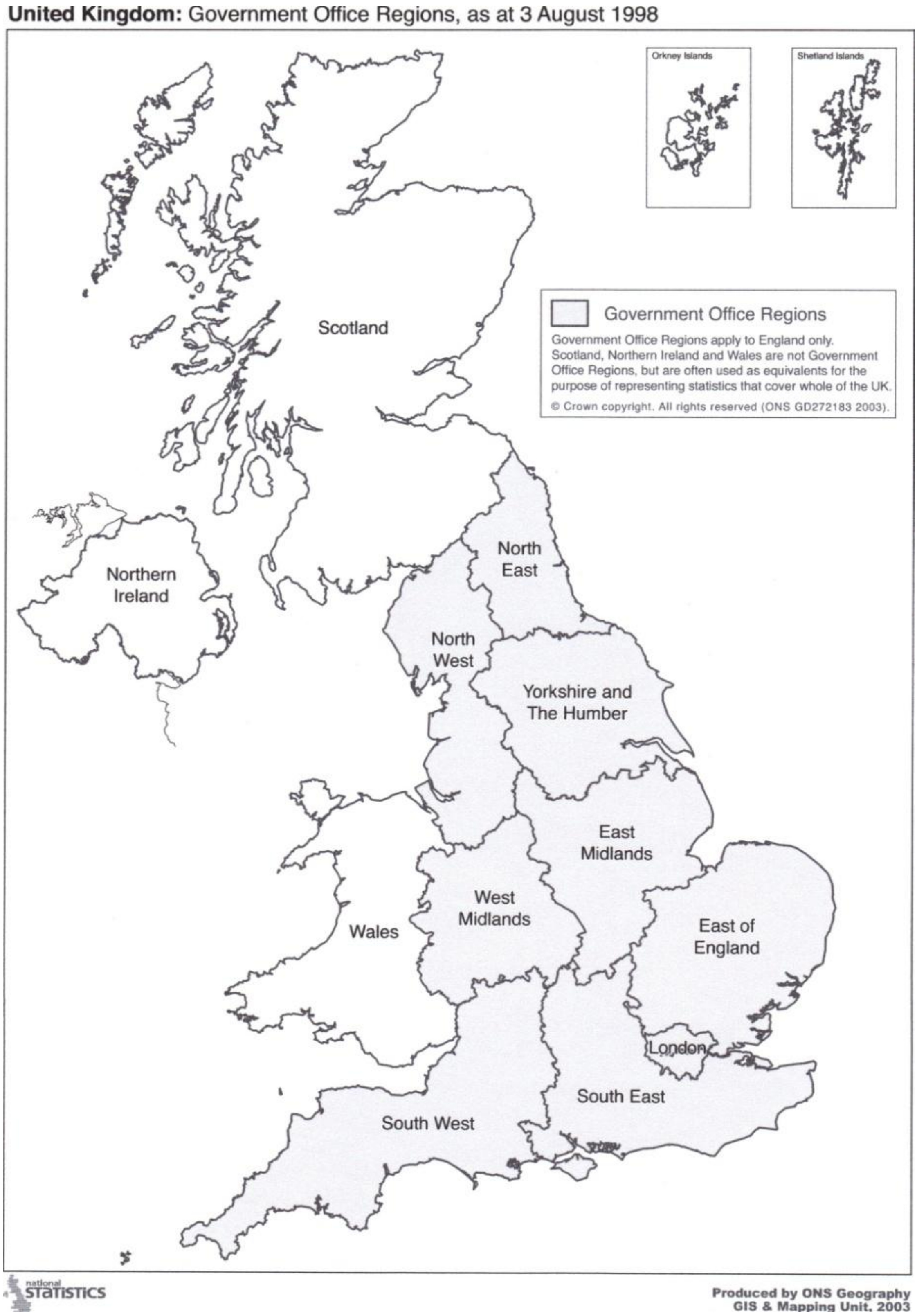


Figure 4.1 Map of Government Office Regions for the period subject to the study

Nelson and Nelson (2003), in their study using data from the United States of America, show that geography based upon state as a category was a significant variable having an effect upon the yields of investment properties. Their study shows that regional diversification can take place either with or without property type diversification. This study shows that diversification by region is an important element in the creation of Markowitz's (1952) *efficient portfolio*. Nelson and Nelson state that the main motivation behind diversification is to eliminate specific risk. They find that whether diversification is based upon region, property type or other factors, either singularly or in combination, it is still no substitute for the detailed analysis of each property. However, they show that regional diversification does afford a useful, additional level in achieving an *efficient portfolio*.

Another study also based on data from the United States of America was conducted by Cheng and Roulac (2007). Cheng and Roulac's study shows that although region is an important factor in determining yield for property type, it affects some types of property to a greater extent than other types. However, as with other researchers, such as Nelson and Nelson (2003), Cheng and Roulac find that detailed analysis of each property by investors and their advisors is needed to ensure the best performance of portfolios.

Enever and Isaac (2002) identify individual companies within an industry-type as having a risk when assembling share investment portfolios. That is to say that there is a risk by company. If individual companies within an industry-type can be considered as posing an investment risk within the context of a share portfolio, then it is logical that such risk shall remain for such individual companies within the context of a property investment portfolio. Moreover, since each specific company may have in general terms leases that are different from the lease of the other companies within the sector, then the associated risks will vary across the sector by company. In particular, the rent review dates and the lease expiry dates may in general terms vary by company. Such variations may well be extenuated where each company within the sector has disposed of tranches of properties on sale-and-leaseback terms at different times. Since the retail banking companies have disposed of tranches of banking-halls on sale-and-leaseback terms at different times, it follows that this will extenuate any risk attributable to individual banking companies.

Time is identified by some of the literature as having an influence on the yields of investment property. MacGregor and Schwann (2003) suggest that sub-markets should fluctuate in tandem over time as a reflection of the macro-economic cycle. Therefore, the findings of their research suggest a correlation of regional property yields both within specific property types and within real estate as a whole. Using a dataset that had been used in previous studies, MacGregor and Schwann reached similar conclusions to those studies. They placed cases into the eleven economic planning regions of Great Britain in the same way that Hoesli *et al.* (1997) had done. The study undertaken by MacGregor and Schwann (2003) found little difference in the yields between each of these eleven economic planning regions. This means that there was limited scope for creating an *efficient portfolio* by diversification across these economic planning regions. However, the findings of MacGregor and Schwann do suggest that property type has a significant influence on yield. Nevertheless, their study does show that each of the main generic property types performs in accordance with a national cycle or cycles. Therefore, there is every reason to expect that banking-halls, as a sub-type of property, would also be subject to cyclical effects. Such cyclical performance is consistent with macro-economic factors and is accepted as being a consequence of such factors (MacGregor and Schwann). The cyclical nature of property yields over time is borne out by the research of Krystalogianni and Tsolacos (2004), which shows the correlation in the movement of yields across a range of investment media, of which property is just one element, and allowing for any time lag.

Baum *et al.* (2006: p76) are quite specific in identifying the factors that influence the yields that investors will seek with respect to any given property investment. Some of these factors appertain more to *market risk*. These are identified as being: the overall level of interest rates in the macro-economy; legislative factors; and prevailing tax liabilities. These are factors which can be expected to alter over time. Therefore, over time these particular factors may respectively exert greater or lesser influence than the others. This is accepted and recognized in the business, financial and economic fields as something that can and does happen. Also, these factors are not always easily quantifiable. For these reasons and for the purpose of better presentation, these factors are commonly grouped together in those fields as one factor: *time*. As a result, it can be argued that these influences can be grouped together as *time* in the analysis of the factors influencing property investment yields. Baum *et al.* also cite liquidity of the investment, volatility and the costs of management. All property is an illiquid form of investment, and the costs of transfer are high compared to most other forms of investment. In the United Kingdom, the purchaser pays Stamp Duty Land Tax, which is levied at three different rates according to the purchase price. Hence, in terms of the Stamp Duty Land Tax, three bands arise with that comprising the cheapest property attracting the lower rate of tax, and that comprising the most expensive property attracting the higher rate of tax. Accordingly, the rates of tax levied are based respectively upon small, medium and large lot sizes.

Specific risk factors identified by Baum *et al.* as having an effect on yield are: the physical attributes of the building; location; and security. They define the physical attributes by both design and the state of repair of the building. Although they cite location, they do not define the level of locality as being most relevant. However, Baum *et al.* (2006: p 246) do show that in a valuation report, a valuer should report on both the very specific location and the more general location, in terms of other towns and cities, of a property. Baum *et al.* define security by the specific tenant, the rent and the terms of the lease.

The specific factors that can be deduced from Baum *et al.* (2006) as influencing property investment yields are:

- Property design and state of repair
- Location at both the very local level and at either the regional or sub-regional level
- The tenant's security
- The rent
- Lease Terms

Furthermore, to the *specific risk* should be added those *market risks* identified by Baum *et al.* (2006), which are all capable of being represented by the factor *Time*.

Research undertaken by Dunse *et al.* (2007) sought to model office yields in the United Kingdom on the basis of major cities. Although their research was focused on offices, it does give some insight into the movement of yields in the United Kingdom for a specific type of property over time. Although they find that the yields of offices are heavily influenced by the presence of the nearest city, which may not necessarily be the case for other types of property, they do establish that yield cycles for a particular city tend to follow national yield trends. Dunse *et al.* argue that an office market within a given British city and its hinterland is dominated by that city subject to some local nuances. They show that yields in any given city do follow national trends and that the two are linked. They also show that fluctuations in office yields follow macro-economic factors as well as stock market indices. The findings of Dunse *et al.* therefore do add weight to the hypothesis that time is a factor that influences yield.

4.9 The factors considered by professional practice.

Traditionally, a literature review underpinning research draws upon the corpus of extant academic literature to develop theory in order to guide research. Some of the leading international property consultancies have their own research departments. It follows that those undertaking research within the context of their own professional practice should also review the appropriate output of the research departments of those leading consultancies where such output exists. The factors identified by the corpus as having an effect on yield should also be considered in the context of current professional practice. As some of the datasets used by the research departments of the leading property consultancies are different, it is to be expected that some of the findings may not necessarily be entirely consistent between those practices. However, generalized findings from professional practice, especially if they are supported by the review of extant academic studies, will help to generate the theory that will direct this research.

A common theme that emerges from the research departments of the leading property consultancies is that property type is a factor that has an influence upon yield. Both academic researchers and those in professional practice assign three principal categories at the first level of property type: retail; office and industrial. However, some researchers further sub-divide these three main property types into sub-property types. The approach of such further sub-division tends to be favoured more by professional practice than in the extant academic on the subject. Notwithstanding such differences in approach, any findings that any kind of division into property type might produce an effect on yield provide justification that property type should be investigated in future research as a factor. Not only does it justify *banking-hall* being studied as a separate sub-type of property, but it means that further thought should be given to additional sub-division of *banking-hall* as a type. An obvious further sub-division is by the actual banking company operating a given bank branch. Certainly, British retail banks have long been categorized by market commentators into two distinct groups: the leading banks, traditionally the *Big Four*; and the secondary retail banks and building societies. This distinction became more obscure with the creation of a fifth large banking group through the takeover of the Bank of Scotland by the former Halifax Building Society. Even within the leading retail banks, there have been differences inasmuch as each has been subject to different performance and risk factors such as differences in profitability, balance sheet strength and market exposure. Coupling these considerations with the fact that other banking companies have always fallen into the secondary banking sector must mean that *banking company* ought to be looked at as a further division of the property sub-type. The current research relies on data up to July, 2006. The credit crunch since mid-2007 has shown that even amongst the main five United Kingdom banking groups, some have been more adversely affected than others by the banking crisis and the recession. Some have been more heavily exposed to adverse risk, with the result that some have had to be part-nationalized and have to do more to re-capitalize (Cushman & Wakefield, 2008). Therefore, in the light of the credit crunch, *banking company* can reasonably be considered as possibly having a greater effect on yield than before the crisis.

Drawing on reports published between 2008 and 2009, property type was confirmed by the research departments of the leading property consultancies based in London as having an effect on yield (Tipping and Lam, 2010; Cushman & Wakefield (2009), Jones Lang LaSalle (2009a, 2009b, 2008a, and 2008b), CB Richard Ellis (2008a and 2008b) and Frank Knight (2009). However, there are some inconsistencies between the findings of the respective research departments. Thus, although Lambert Smith Hampton (2009) did find that property type had an effect on yield, they found in the first quarter of 2009 that office yields had risen more steeply than retail yields, whereas Cushman & Wakefield (2009) found the converse. Jones Lang LaSalle (2008b) found that within the retail property type, there were movements of the yields of each of the sub-property types, including retail bank premises, which moved independently of one another. Hence, retail bank premises reacted independently of other retail categories during 2009 and saw the largest increase in the retail sector during the period (Knight Frank, 2009). According to Knight Frank (2009), this may have been a result of anxieties about the banking sector.

Whereas some of the academic literature, especially that relating to studies based on the United States of America, suggests that region has an effect on yield, evidence from professional practice in the United Kingdom is sketchier in relation to region being a factor

(Tipping and Lam, 2010). Some of the output by the research departments of the leading property consultancies does not produce any results with respect to regions (Tipping and Lam, 2010). Where results are produced showing regions, region is shown to have some slight effects on yield (Cushman & Wakefield (2009); CB Richard Ellis (2008a)). The findings of Cushman & Wakefield (2009) suggest that any effects that are attributable to region are different according to property type. Thus, the yields for the retail sector by region were more closely grouped together than those for other property types. CB Richard Ellis (2008a) similarly found that the effect on yield by region was not consistent across all property types. Cushman & Wakefield (2009) found that those properties in the retail sector nearer to London and the South-east generally achieved lower yields than those in the rest of the country. Likewise, the output of CB Richard Ellis (2008a) suggests that although the effects on yield were not consistent across all property types by region, there remained a north-south divide for retail and office properties (Tipping and Lam, 2010).

Some of the output by the research departments of the leading property consultancies focuses on cities, as opposed to regions, as a geographical factor (Tipping and Lam, 2010). Such output labels cities as *location*. In reporting on Cushman & Wakefield's research into yields in early 2009, Whitmore (2009) suggests that location has an effect in combination with lot size.

In addition to Whitmore (2009) suggesting that lot size has an effect on yield, some of the research departments of the leading property consultancies show likewise (Cushman & Wakefield 2009; Jones Lang LaSalle, 2008b and 2007). According to Whitmore (2009), who does not distinguish retail from other property types, while lot size does have an effect, there is a suggestion that it does so in combination with other factors. Hence, reporting on the research undertaken by Cushman & Wakefield, Whitmore (2009) has suggested the clear emergence of a two-tier market during the early part of 2009 with larger lots in secondary locations seeing larger increases in yield than those property investments that are smaller and in prime locations (Tipping and Lam., 2010). In this context, the effect of lot size may be more complicated, since it may be having an influence in combination with another factor. Although he does not quantify the effect, Whitmore suggests that during the adverse macro-economic conditions of the period, yields of larger properties and those of properties in secondary locations would rise more in relation to the yields of smaller properties and properties in prime locations. Therefore, there is also a suggestion that *Time*, reflecting the economic cycle may be a factor having an effect.

Reports published by CB Richard Ellis (2008a and 2008b), although not giving data on actual lot sizes, do make some interesting suggestions. Whilst these reports do not report upon lot sizes within the retail sector, nor indeed within any other sector, they do differentiate between standard shops and retail warehousing. These reports do show that yield movements of both yield and rental values vary between standard retail and retail warehousing. However, it must be borne in mind that these two property sub-types tend to be found in different locations. Therefore, it is once again important to appreciate that any significant effect by lot size might be operating in combination with another factor: this time location.

Jones Lang LaSalle (2008b) reported that across all three of the three main commercial property types, both rental values fell and yields rose during the preceding months in 2008. It would therefore be logical to expect average lot sizes to have fallen during that period. However, on the contrary, average lot sizes within the retail sector rose during that period (Jones Lang LaSalle, 2008b). Indeed, Jones Lang LaSalle (2008b) show that the average lot size in the retail sector rose by twenty-three per cent between the second and third quarters of 2008. Jones Lang LaSalle (2009b) show that average lot sizes continued to rise during the first quarter of 2009. They show that average lot size with respect to the retail sector rose almost twelve per cent compared with the previous quarter. This would suggest that during that quarter and indeed the immediately preceding quarters, investors were choosing to retain the smaller lot sizes and to dispose of the larger ones. Hence, during the period in question, it would appear that the smaller lot sizes were more sought after by existing investors. Such a conclusion would accord with that which could be drawn from Whitmore.

Although the current research is focused upon British retail bank premises sold before the credit crunch, Cushman & Wakefield (2009) report that following the credit crunch, banks were looking to reduce the space of their operational property where possible. Whilst that report relates to the banks' office space rather than operational retail banking premises, it is indicative that in the light of the banking crisis and the economic downturn, the banks are mindful of the cost of occupying surplus property. The suggestion, therefore, is that physical size of all operational bank premises is a factor influencing the banks' demand for premises. It follows that where investors are aware of this, they may prefer to invest in smaller premises more likely to be demanded by the banks. Since physical size, in combination with other factors, may have an influence on lot size, as defined by capital value, it follows that lot size may well have an influence on yield.

There is very limited output by the research departments of the leading property consultancies on the banking sector and in particular retail bank premises. The research published by Cushman & Wakefield (2009) relates to operational property other than British banking-halls. However, Jones Lang LaSalle (2009b and 2008b) do include within their investment market reports some data appertaining to the yields of retail bank premises in relation to the yields of other sub-types of retail premises in the United Kingdom. In both reports, Jones Lang LaSalle treated banking-halls as a sub-sector of the retail sector. They show that in the third quarter of 2008, the average yields of banking-halls changed less than those of the other analyzed sub-sectors in the twelve months to September, 2008; whereas in the twelve months to March, 2009, banking-halls changed more than most of the other retail sub-sectors. The latter was attributed to market anxieties about the potential consolidation of retail banking. Notwithstanding this, such research output demonstrates that not only the three main property types, but banking halls as a sub-type, can influence yields differently. It also demonstrates that any change in the influence on yields can occur at different rates between sub-types, however defined. Furthermore, it also demonstrates that such rates of change can themselves alter over time. Thus when the market perceives either a beneficial or an adverse risk specific to the retail banking sector, the yields for that sub-type may react differently to the overall market. Such market reactions adapt over time. They can be triggered by a perception of either specific risk or market risk.

4.10 Summary

4.10.1 Main factors influencing yield

The review of the academic literature and the reference books suggest the factors that are most likely to affect property investment yields. A search of the extant literature failed to unearth theory relating to the factors most affecting the property investment yields of British banking-halls. Therefore, theory needed to be deduced from the academic literature relating to the commercial investment property in the generic sense. Where the literature appertaining to commercial investment property is scant with respect to the United Kingdom, it has been necessary to look further afield. Furthermore, where in particular the earlier literature does not produce findings directly with respect to factors affecting yield, it has been necessary to seek direction from another perspective. Such a perspective is how property investment portfolio managers have identified factors to consider in portfolio diversification. That is because in the creation of the *efficient portfolio*, investment managers need to reduce *specific risk* by identifying factors that may increase variance of yield due to those factors having the greatest effect on it (Isaac and Steley, 2000).

Further academic literature and property investment reference books have been used to identify the main factors affecting commercial investment property at the generic level. These sources support the findings and the theory derived from the earlier literature on the diversification of property investment portfolios. Recent output by the research departments of some of the leading property consultancies help to confirm the findings of the academic literature and of the reference books.

The earlier studies into investment property capitalization rates are focused on a stock market approach rather than a property investment approach. They are also based upon datasets from the United States of America rather than from the United Kingdom. However, these earlier studies provide a starting point from which theory can be developed. Furthermore, the corpus of theory has evolved since these earlier studies and some of the more recent studies are based upon datasets from the United Kingdom.

The findings of Ambrose and Nourse (1993) established that property should be considered by property type rather than generically as a single asset class. It is now established practice for researchers not to treat property as a single asset class. At the first level, all retail property is grouped together as a single category (Chen *et al.*, 2004; McGregor and Schwann, 2003; Ambrose and Nourse, 1993).

Geography is shown in the literature to have an effect upon property investment yields. Geography is capable of being measured in different ways. The two obvious main distinctions are between region and location in the narrower, more parochial sense. Modern Portfolio Theory recognizes the importance of property investment diversification by region. Drawing on the approach of treating region as a variable in the United States of America (Cheng and Roulac, 2007; Nelson and Nelson, 2003; Ambrose and Nourse, 1993), region is established as a variable having an effect on yield. Research output by the research departments of the leading property consultancies does rely upon regions based upon the Government Office Regions. However, the findings of such research do not necessarily always identify an effect based upon these regions. Notwithstanding the absence of such a regional effect, other findings have found the existence of a super-regional, or provincial, effect. Whilst some sources do support a super-regional effect based upon a north-south divide by the three main property types (Guy and Henneberry, 2000; Hoesli *et al.*, 1997), there are other sources which do not necessarily support such an hypothesis for the retail sector (Enever and Isaac, 2002).

In their study, Ambrose and Nourse treated the amount outstanding on mortgage advances with respect to each property as a variable influencing return. This does suggest that the value of a property may have an effect on yield. In the United Kingdom, professional practice applies the term *lot size* to reflect the size of a property's capital value. Whitmore (2007) and Baum *et al.* (2006: 76) suggest that *lot size* is a variable that may influence yield.

Property yields do change over time as a reflection of the macro-economic cycle and market risk. Krystalogianni and Tsolacos (2004) demonstrate the cyclical nature of property investment yields. The study by Dunse *et al.* (2007) into the United Kingdom office market also demonstrates the effect of time on the yields of that property type.

Each of the factors identified as having an effect on *yield* may not necessarily be operating in isolation from other variables.

4.10.2 Hypothetical predictive framework for British banking-hall yields

Neither the academic literature nor the property investment references books identify the factors influencing property investment yields specifically with reference to banking-halls.

Banking-halls are a sub-type of commercial property. On this basis, it has been necessary to look at factors influencing investment yields with respect to commercial property. More specifically, the retail type of commercial property, of which banking-halls are a sub-type, has been examined. It can be deduced from theory what are likely to be the principal factors influencing the yields of property as a generic asset class. Beyond this, for the purposes of the current study any theory relating specifically to British banking-halls needs to be added.

Although both academic researchers and those in professional practice now commonly assign retail, office and industrial as the three principal categories of property type at the first level, professional practice often further categorizes property into sub-types at other levels. Moreover, a limited amount of the output by the research departments of the leading property consultancies does treat banking-halls as a separate sub-type of retail property (Jones Lang LaSalle, 2009b and 2008b).

Enever and Isaac (2002) identify individual companies within an industry-type as posing a specific risk when portfolios of company shares are being put together. In which case, it is logical that such individual companies pose a specific risk when portfolios of investment properties are being assembled. On this basis, it is justifiable to extend property sub-type to a further level of the specific banking company tenant when investigating British banking-halls as a property investment.

Accordingly, on the basis of the corpus relating to property investment in general, together with the very limited literature relating to British banking-halls, extant theory suggests that the main factors influencing the property investment yields of British banking-halls are:

- Region
- Tenant banking company
- Lot size
- Time

These four main predictor variables, together with secondary influences also identified in the literature, are derived from the theory as potentially being the variables exerting the greatest influence. Lesham and Trafford (2007) argue that in most forms of either deductive or inductive research, the conceptual framework ought to be placed after the discussion of the theoretical perspectives. The resultant map of the conceptual framework at this stage is reproduced in Figure 4.2. However, it can only be assumed that the factors derived from the theory are the correct ones. Therefore, it is necessary to confirm these factors through the qualitative and quantitative studies.

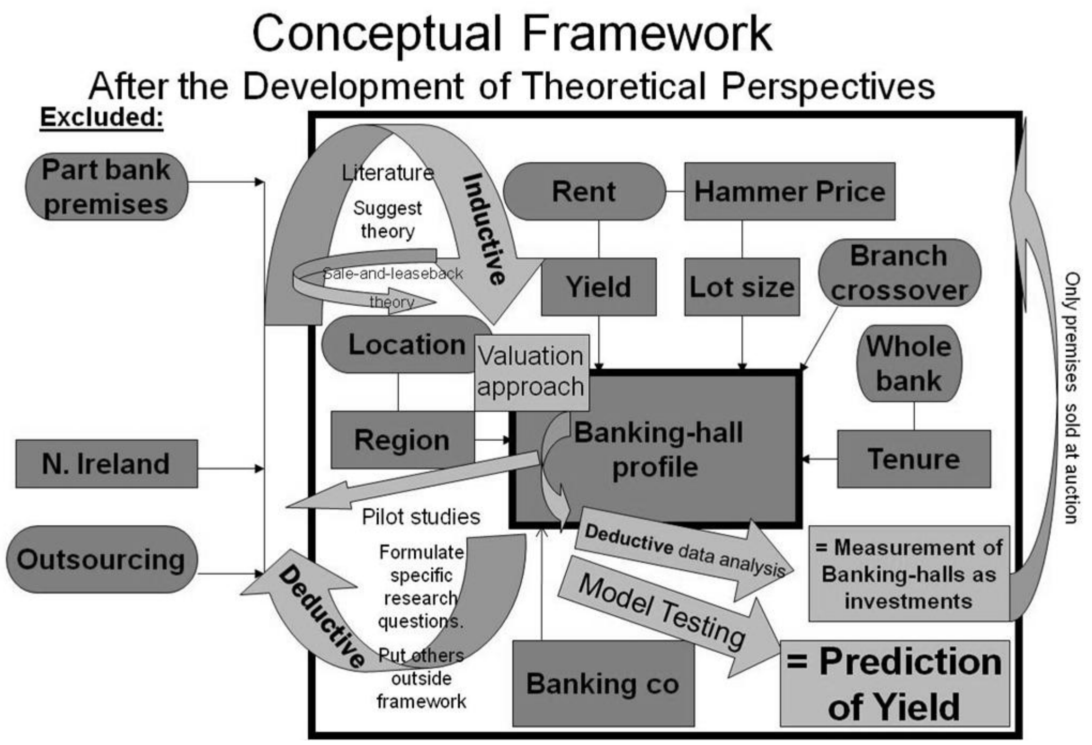


Figure 4.2 **The conceptual Framework after the Development of Theoretical Perspectives**

5 The Research Design and Methodology

5.1 Introduction

Empirical data with respect to banking-halls sold at auction as property investments during the study period exist in the public domain. Once observed and recorded, these data need to be included within a designed research which is robust and which remains valid and reliable. This chapter examines research paradigms in the context of the available data in order to identify the appropriate research methodologies and methods.

The auction data lend themselves to a quantitative study. However, methodological triangulation makes research more robust. The identification of an appropriate qualitative study enables this research to achieve methodological triangulation. Such a qualitative study also confirms the factors and factor levels to be analyzed and tested in the quantitative study. Both of these methodologies are also underpinned and further triangulated by the development of theoretical perspectives from extant literature.

The methods used within the respective qualitative and quantitative studies are identified. Research design should take account of threats to validity (Creswell, 2009, 162-165). Where threats to validity have been identified, measures are taken to isolate such threats. In particular, validity issues arising out of mixed use premises and different types of tenure are considered. Research design also needs to ensure that the research is reliable. Again, where threats to reliability have been identified, steps are taken to address the issue.

Drake and Heath (2011) show that research by the professional practitioner should be undertaken within the conventions and confines of their professional practice whilst also trying to meet the expectations of academic research, especially in relation to robustness. This has been borne in mind.

5.2 The process of investigation

The research design provides the means by which the investigation will be undertaken. It determines how the hypotheses will be proven and therefore how the research question will be answered (Collis & Hussey, 2003). Therefore, it will determine the research methodologies. Collis and Hussey also state that it is necessary to be able to defend the methodologies selected for the investigation and to be able to defend the approaches to data collection and analyses. Vogt argues that research design is both an art and a science itself that is a necessary process in order to obtain valid answers to the research question (Collis & Hussey). According to Bryman (2004), research design is a framework for the investigation.

The framework for this study is founded upon theory being inducted through a review of existing literature and reference to existing professional practice. Continuing review of the literature and reflection results in the refinement of the theory. Hypotheses are developed from the refined theory. At this point, the investigation transforms into deductive research. This deductive research will focus upon the collection, recording and testing of empirical data, which will be used to test the hypotheses. The results of the analysis will be used to answer the research question.

The research framework for this investigation is treated as having three levels. At its most practical level, likely variables identified from professional practice are included. At a more theoretical level, existing practice and knowledge, together with some pilot studies taken from Stage 1 of the research program, are used to develop the theory and to reduce the number of variables to be tested. At the third and higher level, the research relies upon two processes: first it uses the literature to further refine the theory on sale-and-leaseback; and then it applies statistical analyses to the empirical data to test the remaining variables.

During the current interest in the sale-and-leaseback model in the United Kingdom over the last two decades, there has been an increasing amount of literature published on the model. This literature appears in both professional publications and academic journals. It is this literature that is reviewed during the initial, inductive research stage. Some of the literature, especially that published in the professional publications, may be contradictory. Reflection and experience from professional practice are used to arrive at some synthesis in the theory. Furthermore, data from research journals are used in this processing of arriving at the synthesis and in triangulating the data. The synthesis of the inductive research gives greater focus to the variables likely to be tested.

Once the Stage 1 case and pilot studies had confirmed the variables to be tested and the likely outcome of the investigation, likely hypotheses were formulated. These are confirmed by literature reviews undertaken within the main study. Following the formulation of the hypotheses, the research takes on that of a truly deductive nature.

The deductive research relies upon the collection, recording and testing of empirical data. To prove the hypotheses, this research relies upon the recording and testing of data of banking-halls sold at public auction through the London property auction houses. All of the data sets are obtained either directly or indirectly from material published both prior and after the auctions in question. The directly obtained empirical data are those data either previously published by the auctioneers and their agents or recorded by the researcher at the time of auction. Indirectly obtained data are calculated from the directly obtained data.

The testing of the empirical data is undertaken through statistical analyses. The analyses will focus upon the data recorded from the sale of banking-halls sold at public auction through the London property auction houses. From a valuation perspective, the most important variables are the hammer price and the rent reserved. From these two, a third vital variable can be calculated: the yield. However, the yield needs to be clearly and precisely defined. The term *yield* is capable of many definitions. In order to achieve validity, it is necessary that the calculation of the yield should be based upon the appropriate definition. Therefore, investigation into what constitutes the appropriate definition of yield is undertaken.

Other relevant and important variables are also applied having regard to the practicalities and limitations of this research. Dates, regions, lot size and tenure are all variables for which data are collected. Branch crossover is another variable that ought to be tested. Equally, relevant is whether a property is entirely a banking-hall or is part let to another party using it for non-banking purposes. Not all possible variables are capable of analysis within the practicalities and limitations of this research. Therefore, through the refinement of this research, some variables are inevitably omitted. Where this occurs, it is recommended, where appropriate, that these issues are investigated by subsequent research.

The statistical analyses are mainly undertaken through the application of the 'R' statistical package computer program. Before data sets are input into the program, all the empirical data are hand written into data collection sheets. The data sets are also recorded in a field book. Each data set comprises the empirical data relating to any one property sold at auction. Where any one banking-hall is sold more than once, it is treated for the purposes of the statistical analyses as a separate property. The justification for this is that on re-sale some of the data with respect to each banking-hall will change. The date will certainly change. The data with respect to the hammer price and the rent reserved, and therefore the yield, will almost certainly change. Since the market can be expected to change over time, the data sets on re-sold properties should at any given time reflect that change.

After being transcribed into the 'R' computer program, the data sets are checked for typographical errors. Corrections are made as appropriate. Following the pilot studies, further decisions are made as to how the empirical data should be analyzed and presented. For example, decisions are made about the most appropriate way of banding the data. In particular, the variable with respect to sale date is refined both for practicality and to reflect market practice. The London property auction houses do not conduct their sales in accordance with a regular time-scale. The auctions are held on random dates. The consequence is that date variable records a large number of random dates. These would neither be easy to analyze nor present. Furthermore, the market and professional practice tends to analyze and present the auction data on a quarterly basis. Therefore, an additional variable, converting actual dates into quarters, is created. It is the quarter variable, rather than the date variable, that is used for analysis and presentation.

5.3 Sources of the data

A large sample of empirical data based upon the rents reserved and the hammer prices of British banking-halls sold at auction between 1997 and July, 2006 exists. It was observed and recorded. This large sample of data comprised almost all of the population for the study period. Such dataset forms the principal means of proving or disproving the hypotheses.

The main research question is:

“How can property investors select freehold British banking-halls that are likely to provide the highest initial yield on their investment?”

Following on from this, the following hypotheses are tested:

- There is some regional disparity in yields
- Lot size does make a difference inasmuch as there is increased demand for smaller lots
- Tenant banking company has an effect
- Time has an effect
- The appropriate Investment Property Databank (IPD) index has an effect.

Time is tested as being representative of the macro-economic cycle. Since professional practice relies on indices published by the Investment Property Databank (IPD), further analyses and tests are undertaken substituting the IPD United Kingdom Retail Property Index for time.

The empirical data is valid insofar as it has been accurately recorded. In particular, the data with respect to rents reserved and hammer prices, which are pivotal to the research, are accepted as being accurate. Unless they are truly accurate the property investment market would not be able to function properly. Both the auction houses and third parties relying on the data need to have confidence that the data are entirely accurate. In addition, much of the data were triangulated by having been independently observed and recorded by the researcher. Such independent observation was made either by being present at the auctions concerned or by being connected to those auctions by internet link.

Not all of the empirical data were necessarily reliable. There were some outliers. Such outliers needed to be checked for accuracy. Also, other facets had to be checked and verified. For example, the way in which the region of a property had been recorded at factor level needed to be checked and then corrected in those very few cases where wrongly recorded by the auction houses. Each auction might produce a slightly different result for inexplicable reasons. A particular investor might be prepared to pay more than what would otherwise be the true market value for a specific property. However, it is anticipated that large sample sizes from a large population will render the empirical data reliable in the majority of instances. Even after some data sets were omitted due to being incomplete or the lots not being wholly retail banking premises, over one thousand data sets remained. A proportion of these had to be removed from analysis on the basis of not having a freehold tenure. The necessity for data collapse would eventually leave 691 cases in the dataset for analysis and the building of models to be available in the predictive framework.

The quality of other data obtained during the initial inductive research was more varied. It came from reviewed literature. Much of these data were obtained from professional periodicals. In particular, the professional journals relied upon were *Estates Gazette* and *Property Week*. *RICS Business* was also consulted for material. The consultation of such professional journals is the norm in professional practice.

At the other end of the scale, inductive research reviews peer reviewed refereed journals. Academic papers published in peer reviewed refereed journals and at conferences are reviewed. These have the greater credence for having been peer reviewed. These peer reviewed papers give much insight into the separation of operational property from core business activity in a generic sense. Moreover, they give more specific insight into the theory relating to sale-and-leaseback as a model for the holding of operational property.

Academic papers do provide data and theories on sale-and-leaseback as a generic model. However, they are insufficient to provide data on the adoption of the model by the British retail banks for the holding of their operational property. In order to obtain more insight into the British retail bank's adoption of the model, it is necessary to refer to the professional journals. These professional journals also give greater insight into how the investment market treats British banking-halls as media for property investment.

5.4 Emergent theories from the Stage 1 case and pilot studies

Grounded theory is the research design that most lacks a theoretical framework (Leedy and Ormrod, 2010). However, Drake and Heath (2011: 103) argue that grounded theory has a place in allowing theoretical inferences to be made from the data. They argue that grounded theory also allows for research processes to take account of reflection from within the context of professional practice (Drake and Heath, 2011: 45). Grounded theory is capable of being used in qualitative study to provide abstract theories and interactions grounded in the respondents (Creswell, 2009: 13).

The grounded theory approach assumes that theory will emerge from analysis of the data (Leedy and Ormrod, 2010: 142). Leedy and Ormrod argue that instead of relying upon extant research literature, grounded theory allows theory to emerge from the researcher's reflection and interaction with the data. They show that grounded theory is an iterative process, which relies upon the researcher collecting data, looking for emerging themes from those data, collecting more data and refining theory arising from analyses. Thus, emerging theory is used to explain how certain interactions may influence a dependent variable.

Although grounded theory is more commonly associated with phenomenological paradigms, especially in the social sciences (Leedy and Ormrod, 2010: 142; Collis and Hussey, 2003: 60 and 73), there is no reason why it should not be used outside a purely phenomenological paradigm. Whereas a positivistic paradigm assumes that theory will be developed inductively resulting in hypotheses prior to the study, grounded theory actually allows theory to grow out of the observed data (Collis and Hussey, 2003: 73). Leedy and Ormrod (2010: 146) show that grounded theory may be used to evolve theory within a paradigm that has some positivistic attributes in which it is used in research within the context of professional practice.

Case studies followed by pilot studies, based upon two datasets of Midland Bank and Bradford & Bingley premises sold at auction on sale-and-leaseback terms, were undertaken during Stage 1 of the Professional Doctorate programme. In this way, emergent theory was actually evolved from those data in the manner of grounded theory. At that point, possible hypotheses were identified for testing. These were as follows:

- The less reversionary banking-halls are in especially high demand from investors;
- There is some regional disparity in yields;
- Lot size does make a difference inasmuch as there is increased demand for smaller lots;
- Premises let to the smaller banks and demutualized building societies do not necessarily command lower yields than those let to the main retail banks;
- The influx of new, amateur investors has pushed up values; and
- Branch crossover has resulted in closures through rationalization.

The proviso was made that the testing of such possible hypotheses would be subject to there being sufficient data with respect to each. Hence, it was suggested, for example, that due to insufficient data, it would probably not be possible to test within a positivistic paradigm that branch crossover had resulted in closures due to rationalization.

The case studies from Stage 1 of the Professional Doctorate, based upon grounded theory, were used to suggest likely hypotheses for further investigation in the present study. Notwithstanding that grounded theory had been used in Stage 1 of the Professional Doctorate programme, a new research design was formulated for the current study. Once the present study was commenced with the benefit of emerging theory from Stage 1, greater robustness was given to the current research through a full review of the literature and the development of theoretical perspectives before qualitative and quantitative studies were undertaken. It was further concluded that the emerging themes from Stage 1 needed to be tested and triangulated through more extensive research. It was found these emergent themes only resulted in preliminary conclusions. The research design of the current investigation needs to be such that it is robust and appropriate in its own right, having only drawn upon the emergent theories from Stage 1 as an initial guide for those hypotheses to be tested.

5.5 Designing a framework for the current study

Research design formulates a framework which directs the entire process of the study. The built environment is an area that transcends a number of fields including the natural sciences, social sciences, engineering and management (Amaratunga *et al.*, 2002). Amaratunga *et al.* show that it is imperative to utilize the appropriate methodologies and techniques for research in the built environment to remain robust. They argue that due to built environment research having in the past been criticized for having a more narrative, rather than an empiricist approach, when constructing reality, the research design should address such a shortcoming. The difficulty in applying social science methodologies is that the social sciences normally adopt a design research that reflects a phenomenological paradigm. That is because most of the research within the social sciences is focused upon qualitative study. However, the final design of a research project is influenced by the nature of the problem that the researcher is trying to solve (Leedy and Ormrod, 2010: 4).

The current research is into identifying ways in which investors and professional practitioners may create a toolkit or predictive framework to be used in building an investment portfolio of retail bank premises that maximizes initial yield. As a result, much of the data will be empirical in nature. A very large part of the data will comprise a dataset of capital realizations, rents reserved and yields relating to British banking-halls sold during the study period. Moreover, such dataset will comprise nearly all of the banking-halls sold between 1997 and July, 2006. Therefore, it is not possible to create a designed experiment for this research.

In the absence of a designed experiment, a suitable and robust alternative research design needs to be identified. Creswell (2009: 12) explains how positivism and post-positivism cope with this. Hence, where a truly experimental approach will incorporate separate elements each respectively placed in the treatment or control groups, there are forms of experimental research that do not have such separation. Correlational study is an acceptable form of experimental research with a quantitative framework (Creswell, 2009:12).

Leedy and Ormord show how quantitative research might be undertaken in the absence of a designed experiment. They describe such research as *descriptive quantitative research* (Leedy and Ormrod, 2010: 182). They explain that such an approach delving into observed data can produce real insight into the subject under investigation. Leedy and Ormord further categorize what they describe as *descriptive quantitative research*. One of these categories is correlational research. Leedy and Ormrod show correlational research to be a method of study whereby a difference in a dependent variable is shown to be related to differences in one or more independent variables.

Since the quantitative study in this research examines how the dependent variable, *Yield*, is influenced by independent variables for which there is empirical data from auction sales, it relies heavily upon correlational research.

5.5.1 A research design comprising four levels

Collis and Hussey (2003: 355) define research design as a detailed plan for undertaking research. They add that the design should show it will solve the research problem, should detail an appropriate methodology and the methods used to collect and analyze the data (Collis and Hussey, 2003: 132).

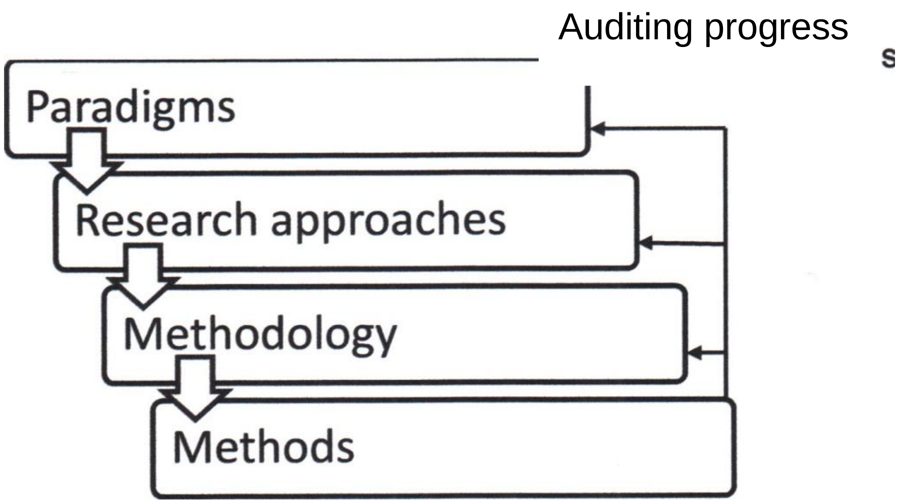
Bryman (2004) defines research design as:

> "A framework for the collection and analysis of data".

He adds that such a framework is determined by objectives such as causality and generalization. These two objectives mean seeking causal effects rather than mere associations between variables and seeking external validity (Bryman).

The research design provides a framework and direction to the research. Before the design can be undertaken, the research paradigm needs to be determined (Collis and Hussey, 2003:

113). Trafford and Leshem (2008) go even further. They show that research design needs to be underpinned by four levels in sequential order. These are shown in Figure 5.1.



Source: Trafford and Lesham (2008) p. 94

Figure 5.1 Levels of thinking about research

5.5.2 Ontology and epistomology

Ontology is about the nature of reality and how it should be studied. It relates to whether the phenomena being studied have a reality external to the researcher or are something that exists in the researcher's mind. It is a divide between the objectivist and the subjectivist. Hence, the objectivist position is concerned with whether reality is external to the mind. This is an ontological view that social phenomena and their interpretations are independent of social actors (Bryman, 2004: 541). In contrast, the subjectivist, or nominalist, position considers subjective reality internal to the mind. Therefore, the subjectivist position is that social phenomena and their interpretations are connected to social actors. Subjectivism is sometimes known as nominalism or phenomenalism. These two opposing positions as to whether reality is external or internal to the mind represent two different paradigms. The convention in the research community is that research based upon one paradigm allows a richer and fuller picture to be formed. However, there is an opposing view that mixed methodologies provide extra data.

Bryman argues that the objectivist position is rooted in positivism and is based upon the natural sciences. Although the objectivist position has its place in scientific enquiry, it has its shortcomings in social research. Bryman highlights realism as a way in which such shortcomings can be addressed for qualitative study. Realism takes neither an objectivist nor a subjectivist position. Instead, in the social sciences, realism has evolved to view social phenomena as having an objective existence, but seen through their effects rather than through empirically recorded observations (Bryman, 2004).

Epistomology is about the study of knowledge. It concerns theories about the sources and nature of knowledge. Therefore, it concerns itself with the methods used for collecting data. It concerns what the research community accepts as valid knowledge of the world. It questions how legitimate knowledge can be gained and, therefore, discriminates between what is valid and what is not. The discussion between positivism and phenomenological study arises again.

5.5.3 The selection of the paradigm

Research is guided and directed by what Creswell (2009:6) describes as the researcher's *worldview*. Creswell explains that this *worldview* has been ascribed different terms by other authorities. These other terms include *paradigm, epistemology and ontology,* and *broadly conceived research methodologies* (Creswell, 2009:6).

According to Collis and Hussey (2003: 48-51), researchers rely upon two principal paradigms. These paradigms are philosophies which provide the framework for research. Fellows & Lui (2003) summarize paradigms as follows:

> "They operate to determine not only what views are adopted, but also the approach to questioning and discovery."

The two main paradigms are normally respectively based upon quantitative and qualitative assumptions (Collis and Hussey, 2003:48). Researchers tend to use alternative labels for these philosophies. More generally, the terms used for these two alternative philosophies are *positivistic* and *phenomenological*. However, Bryman (2004) cautions whether quantitative and qualitative research are necessarily paradigms. He argues that there are aspects of research which are common to each and that neither predominates in the social sciences.

In spite of there being much argument by those such as Kuhn (1970) that research should be conducted within the context of one paradigm (Bryman), Collis and Hussey show that it may be appropriate to transcend more than one paradigm. Using a continuum of core ontological assumptions derived from Morgan and Smircich (1980: 492), Collis and Hussey demonstrate that in the social sciences it is sometimes useful to adopt a paradigmatic position lying somewhere between the positivistic and phenomenological paradigms.

Positivistic and phenomenological paradigms are pure or extreme philosophies. The positivistic paradigm is an objective one which is capable of measurement and which is deductive. In contrast, the phenomenological paradigm is one where reality is subjective and inductive. With respect to the social sciences, Morgan and Smircich (1980) place these extreme points of the two paradigms on a continuum on which they show intermediate positions (Collis & Hussey, 2003). On these intermediate points are placed research methods which adopt varying degrees of each of the two main paradigms.

The current research relies heavily upon the analyses of the auction data of banking-halls sold during the study period. These data are clearly empiricist and positivistic. However, the current study may draw on other data from the experts in the field, who are the auctioneers and their professional staff. The latter data may slant either towards a positivistic or a phenomenological stance. This depends on the methods employed. Clearly, if the data were collected from focus groups it would be more phenomenological. However, such an approach might provide an opportunity for bias to appear (Bryman, 2004, 345-360). Surveys are seen as being positivistic. This is especially so if those surveys are in the form of questionnaires and more so if such questionnaires allow the data to be numbered. Therefore, in order to ensure the sort of empiricism demanded by Amaratunga *et al.* (2002) within a built environment study, the data collected from the responses of the auctioneers and their professional staff are observed through a questionnaire. The form of the questionnaire is framed partly through the literature review on theoretical perspectives and partly through the interviewing of the heads of the auction houses. Interviews may be either positivistic or phenomenological (Collis and Hussey, 2003: 167). To retain a positivistic dimension the interviews ought to be based upon closed questions. Nonetheless, the heads of the auction houses, as the very leading experts, must have some freedom to elaborate their responses. Providing unstructured interviews might allow bias to creep in. It is also a method that is strongly associated with a phenomenological paradigm (Bryman, 2003). Therefore, to reduce the risk of bias and to dilute any phenomenological influence, the interviews were semi-structured. Kane and O'Reilly-De-Brun (2001: 154) state that semi-structured interviews are an appropriate means of triangulating data obtained using other methods.

Bryman reviews the arguments against a multi-strategy research design. He shows that those who argue in favour of strict paradigmic demarcation do so on the grounds of adherence to epistomologic convention and on the premise that quantitative and qualitative study represent separate paradigms. Bryman counters this on the basis that quantitative and qualitative methodologies can be used to complement one another within the same paradigm. He argues that some forms of qualitative study can have what he describes as *empiricist overtones* (Bryman, 2004: 439) in a research project to give a view to reality. Hence, he shows how in the social sciences, qualitative study may be used for the kinds of research that would normally rely upon the scientific methods of the natural sciences (Bryman, 2004: 439). Creswell (2009: 3-18) also shows that a multi-strategy research design may be used. He states that there is a third research design and refers to this as *mixed methods research*. This is sometimes known as a *pragmatic* worldview. The current study into the investment yields of retail bank premises falls within the social sciences. Hence, it follows that the appropriate qualitative methods may be used within a paradigmic position that is slanted towards positivism.

Due to the need to analyze the empirical data based upon the sale of banking-halls by auction, this research adopts a paradigm closer to the positivistic end of the Morgan and Smircich continuum. The collection and analysis of the auction data of banking-halls sold on sale-and-leaseback is scientific and capable of measurement. Therefore, this is very close to a purely positivistic paradigm. However, the development of theoretical perspectives and the qualitative study adopt essentially phenomenological characteristics.

The main analyses and model-building in this study is firmly based upon the data relating to the hammer prices and rents reserved of British banking-halls sold at auction. Comprising almost all of the population, the sample size is very large. The auction data are empiricist and positivistic. So is the questionnaire. The interview element of the survey has the potential to be phenomenological, but that potential is diminished by the used of semi-structured interviews. Although the data have attributes that fall into both the positivistic and phenomenological paradigms, those attributes are shown by the criteria set by Hussey and Collis (2003: 55) to be clearly at the positivistic end of the Morgan and Smircich continuum. Therefore, the current study adopts a paradigm that is much slanted towards a positivistic paradigm.

As a result of adopting a paradigm that is highly influenced by positivism, this research focuses on testing theory. Therefore, it is mainly deductive. However, the positivistic paradigm is tempered by the need to develop the theoretical perspectives that underpin both the qualitative and quantitative studies. It also embraces some qualitative aspects, from which it could be argued, according to Creswell's reasoning, that a pragmatist worldview arises.

5.5.4 Methodology

Methodology is the study of methods. It is about the philosophical assumptions, whereas method is about the actual implementation or technical approach. The framework for the methodology is prescribed by the paradigmic assumptions. Numerous methodologies exist, but ones consistent with the paradigm adopted must be used (Collis and Hussey, 2003: 60). Whilst keeping this in mind, it is also necessary to achieve methodological triangulation through the use of both qualitative and quantitative methods.

Due to the reliance upon extant data of banking-halls sold at auction, it has not been possible to construct a designed experiment. An experiment not only requires an experimental group, but also a control group (Bryman). The data of banking-halls sold at auction comprises numerous cases having numeric and categorical data held in existing data-bases. Such data in this form are ideally suited for a cross-sectional research design (Bryman, 2004: 41-46). The data for cross-sectional research are collected just once over a short period of time (Collis and Hussey (2003: 61). They are commonly used for economic studies and can then be used for statistical analyses to establish correlation (Collis and Hussey). Although cross-sectional studies do not explain why a correlation exists, the correlations can be used for model-building. The current research seeks to provide a framework for predicting which retail bank premises are likely to produce a higher initial yield. That being the case, it is not important to know why a correlation exists. What is important is the creation of a predictive framework for the identification of those premises likely to produce a higher yield. Therefore, a methodology based upon a cross-sectional study is justified.

To ensure robustness of the research, the cross-sectional study needs to be triangulated. The method of triangulation should be one that is appropriate to a paradigm influenced by positivism. Bryman shows that qualitative study may fall within the positivistic paradigm. Moreover, he states that qualitative study may have *empiricist overtones* and may act as a *springboard* to positivist research within the social sciences (Bryman, 2004: 439). On this basis, it is justified to triangulate the cross-sectional study by using an appropriate survey of the professionals regularly dealing with the sale of banking-hall investments by auction. Moreover, such a survey is used to confirm the hypotheses to be tested by statistical analyses following the cross-sectional study.

Robustness of research is also enhanced through engagement with the literature. This research engages with the literature in two ways. Firstly, a literature review is undertaken to provide a deep and broad understanding both about retail bank premises in Great Britain and about sale-and-leaseback as a model for holding operational property. Secondly, a further form of literature engagement is required to provide theoretical underpinning to both the qualitative and quantitative studies. This is especially important in view of the dearth of extant literature relating specifically to banking-halls as a form of property investment.

Further data for fifteen months immediately after the study period is capable of being used to further validate the quantitative study. Such further validation enhances robustness to an even greater extent.

The sequence of methodology can be summarized as follows:

- Literature review
- Development of theoretical perspectives

- Qualitative study engaging professional experts
- Cross-sectional quantitative study
- Further validation of the quantitative study using new data

5.5.5 The methods and techniques deployed: triangulation

An introductory chapter is provided to provide a background to the topic and purpose of the study. This introduces the research objective and suggests hypotheses.

A second chapter outlines the ways in which operational property may be held. This also discusses sale-and-leaseback as a model for holding operational property.

Factors influencing investment in retail bank premises are examined in a further literature review in a third chapter. Together, these first three chapters provide an overall literature review that gives a wide-ranging insight into British banking-halls as investment media.

The chapter relating to theoretical perspectives provides a theoretical under-pinning to the qualitative and quantitative studies. Such under-pinning confirms the hypotheses to be tested. Since there is very little extant literature about retail bank premises as property investments, other sources have had to be used to build the theoretical perspectives. Therefore, the nearest equivalents to banking-halls had to be examined. Since banking-halls were identified as being a sub-type of retail property, the retail property type was relied upon as being the nearest equivalent. Accordingly, the chapter on theoretical perspectives focused on an examination of the factors most influencing the yields of retail property. Notwithstanding that banking-halls are a sub-type of retail property, this approach gives a degree of triangulation of theories.

The qualitative study was purposely designed as one that was compatible with a positivistic paradigm. It took the form of two parts. Each part was to involve the collection of data from those partners and professional staff within the specialist auction houses willing and able to participate. They were in the form of survey research. Surveys, whether in the form of interviews or questionnaires, fall into a positivistic paradigm (Collis and Hussey, 2003: 60 and 66). Leedy and Ormrod (2010:187) state that although some researchers refer to all descriptive, quantitative study as being survey research, they prefer to confine its definition. They contend that survey research is about asking questions of respondents about their experiences, responses and thoughts before recording and tabulating that data (Leedy and Ormrod, 2010:187). Such a method is sometimes called *descriptive survey* or *normative survey* and has an uncomplicated design (Leedy and Ormrod). However, although descriptive surveys have a simple design, it is important to ensure that the personal prejudices of the respondents, or their lack of total frankness, does not compromise validity (Leedy and Ormrod, 2010: 188; Bryman, 2004: 167-168).

The design of this research relies upon the qualitative study being divided into two separate and different forms of survey. All the respondents within the qualitative study were experts from professional practice. Judgmental sampling selects respondents on the basis of expertise and experience in the field (Collis and Hussey, 2003: 158). The two surveys in the qualitative study were respectively based upon interviews and questionnaires. These interviews and questionnaires were a means of drawing on the professional practitioners' expertise and experience. Methodological triangulation is achieved by combining qualitative and quantitative study (Collis and Hussey, 2003: 78). Huberman and Miles (2002) argue that by using more than one form of qualitative study, even greater triangulation can be given to the research. Although, the questionnaire is relied upon to confirm the hypotheses to be tested in the quantitative study and to provide methodological triangulation, it in turn needs such underpinning. Therefore, the design provides for an interview survey of three individuals from professional practice identified as being the leading practitioners in the field. This way, greater confidence is provided for the robustness of the research.

Once the hypotheses to be tested have been confirmed by the qualitative study, the empirical data of the banking-halls sold at auction during the study period can then be analyzed and models built and tested. A quantitative study is capable of giving generalisation from the findings and of providing objectivity (Amaratunga *et al.*, 2002: 20-22). The design for the quantitative study is a cross-sectional one based upon correlational analysis. The cross-sectional study is founded on data in the public domain for the study period and collected over a short period. Those data are then subjected to correlational analysis. The type of correlational analysis undertaken is determined by the type of data and the way in which it is distributed. That is to say that the methods employed in the qualitative study are determined by the way that the variables are arranged into discrete or continuous data.

The principal analyses of the current research are founded upon the observed and recorded auction data for the period. The recording is based upon each banking-hall being placed into a separate case in the dataset and being assigned separate factors to represent each variable. In turn, each category is assigned its own factor level which is recorded with the appropriate factor within the dataset. On that basis, the statistical analyses and tests of the data are repeatable.

Prior to any analyses of the auction data being undertaken, the data were cleaned to correct errors and to remove cases with missing data. The literature on statistics then was used to determine the appropriate statistical methods to adopt for the analyses. This was necessary to determine how to treat categorical data and how to deal with any issues arising from any distribution of the data that had the potential to compromise the analyses and tests.

Much of the available data with respect to banking-hall investment properties are categorical. Field (2005) identifies logistic regression and analysis of variance (ANOVA) as the appropriate methods for analyzing and testing such categorical data.

The availability of new data for fifteen months immediately following the study period permits further validation of the quantitative study. Such further validation enhances the robustness of the research.

The outline of the research design methods can be summarized as follows in sequential order:

- Review of the literature
- The building of theoretical perspectives
- Qualitative Study divided into two parts:
    - Semi-structured interview
    - Questionnaire
- Quantitative study founded on cross-sectional research
- Further validation from post-study period data
- Findings and conclusions

5.6 Validity issues

Issues affecting validity need to be identified and addressed (Creswell, 2009. 164). This can be made easier and more effective if such threats to validity are identified at the outset and the research design framed accordingly. Bryman (2004: 545) and Collis and Hussey (2003: 58-59) show that validity appertains to the soundness of research findings.

Collis and Hussey cite Coolican (1992: 35) who defines validity in the following way:

> 'An effect or test is valid if it demonstrates or measures what the researcher thinks or claims it does.'

Careful consideration needs to be given as to which independent variables should be analyzed and tested. The design achieves this at several levels. Firstly, a raft of theoretical perspectives on factors most influencing the yields of retail bank premises is built. In the absence of such literature specific to British banking-halls, the nearest appropriate literature is reviewed. Since banking-halls are a sub-type of retail property, the raft of theoretical perspectives is built upon a review of factors most influencing the yields of retail properties. This is what Collis and Hussey (2003: 78) describe as triangulation of theories. In providing the raft of theoretical perspectives, the research design further underpins the other methodologies. It also provides methodological triangulation.

At a second level, having been suggested by the theoretical perspectives, the independent variables identified as being likely to affect the yields of retail bank premises can be tested qualitatively through survey. The design allows for the qualitative study to be to be a survey split into two parts. The first part of the survey takes the form of a semi-structured interview to check the theoretical perspectives and to confirm the form of a subsequent questionnaire. The questionnaire provides the findings to the qualitative study, which further underpins the subsequent quantitative study. The qualitative study gives further methodological triangulation to the quantitative study, thereby making the research much more robust.

The main part of the survey was a questionnaire which was designed in a way to limit risks to validity. A questionnaire is seen as being positivistic when comprising closed questions

(Collis and Hussey, 2003: 173). Hence, a closed question questionnaire was adopted. This was conducted on-line individually so that respondents were not influenced in their responses by other respondents. Also to ensure the validity of the questionnaire, the request of the head of one auction house to answer on behalf of all his colleagues was refused. Therefore, the professional experts from that auction house were not included in the questionnaire so as not to compromise internal validity.

Tenure has been identified at the outset as a possible threat to validity. The validity issues appertaining to tenure are twofold. Firstly, the sponsors of this research have made it clear that they are only interested in having the research confined to freehold properties. Tenure has been shown to have an impact on yield (Isaac and Steley, 2000: 25-30 and 126). Had the sponsors not insisted on the research being confined to freehold properties, a *Tenure* factor could have been incorporated with different tenures reflected in factor levels. Secondly, without the need to have a *Tenure* factor, thought still needs to be given as to what constitutes freehold. That is obvious in England and Wales. It is not so obvious in Scotland, which has a different system of tenures. Threats to internal validity can arise where the wrong treatments or wrong data are used in the experimentation and analyses (Creswell, 2009: 162). Those threats posed to validity by tenure are discussed in sections 5.9 and 5.10 below.

Mixed use premises were identified as another possible threat to validity. Hence, those premises that included a residential element or some other non-retail bank use were omitted for the dataset. Only those premises that were entirely banking-halls were retained.

Hence, the research design addresses identified potential threats to validity.

5.7 Reliability issues

Creswell, 2009: 190-191) highlights the need to address matters threatening reliability. Reliability concerns itself with consistency and repeatability. This is more important in the case of positivistic research and especially in the case of designed experiments. Thus, the results should remain very much the same if the research and any testing were repeated

(Collis and Hussy, 2003: 58). Bryman (2204: 543) defines reliability by the stability of the measure of an effect. Within the context of a phenomenological paradigm, reliability is not so important in terms of the stability of the measure of an effect. However, undertaking this study in the context of a paradigm that is in the main positivistic and relying on much empirical data, it is important that the design of this research withstands threats to reliability.

Creswell identifies steps to be taken, where appropriate, to safeguard reliability. Those relevant to this study are:

- Checks to avoid mistakes in the observation and recording of data
- Ensuring that codes applied to data are not corrupted during the research process

The present research addresses the threats identified through Creswell in both the qualitative and quantitative studies.

Hence, in the qualitative survey, the following steps were taken to protect reliability:

- The semi-structured interviews with the heads of the three leading auction houses were recorded in writing at the time
- The questionnaire was conducted on-line using Survey Monkey software, which avoided the use of data coding

During the quantitative cross-sectional study, extra care needed to be taken to protect reliability. Not only did the researcher have to collect and record data in the public domain, but that data had already been put into the public domain by third parties. Therefore, it became necessary to scrutinize each individual datum in each individual case for errors that may have been input by either those third parties or subsequently by the researcher. Also, steps needed to be taken to safeguard against the incorrect coding of data prior to analyses and testing. Accordingly, the following steps were taken to protect reliability during the quantitative study:

- Obvious mistakes by third parties were corrected before the data were entered onto data collection sheets and subsequently into a field book
- Statistical outputs were generated twice, using different statistical packages, to identify obvious outliers which were then double-checked for accuracy
- Since the versions of SPSS (Statistical Package for the Social Sciences) available relied upon the coding of categorical data, the main statistical analyses were subsequently undertaken using the R package which did not rely upon coding, because it was able to read categorical data

Thus, the research design took appropriate steps to safeguard reliability.

5.8 Ethical considerations

Ethical considerations need to be undertaken when conducting research (Collis and Hussey, 2003: 37-40). Effects of the research that may cause harm should be subjected to ethical considerations (Holmes *et al.*, 2006: 292-293). Guidelines for ethics vary between disciplines and institutions. From Collis and Hussey, the main areas for ethical consideration are:

- Illegality or immorality discovered about third parties participating in the research
- Respecting undertakings of anonymity and confidentiality
- Informed consent of the participants
- Dignity of participants
- Integrity and the absence of falsification

The quantitative research relied upon data already in the public domain. The current study relied upon observing, recording and analyzing those data. Therefore, ethical considerations did not arise in the quantitative study. The main analyses relied on the data of premises sold by three banks, which were fully in the public domain and available on the internet. However, data of premises sold by one of the secondary retail banks were observed and analyzed in the preliminary data mining. The data relating to the sales by that bank had been blanked out on the internet source. However, those data had already been placed in the public domain by having been published in *Estates Gazette*. Therefore, no breach of ethical considerations was made by collecting the data from copies of *Estates Gazette*.

The qualitative study relied upon the responses of respondents who were experts in the field. Therefore, ethical considerations did apply in the qualitative study, but to a lesser extent than if the respondents had not been professionals and experts in the field. Before the qualitative study was commenced, the heads of the participating auction houses were contacted by e-mail. They were informed of the context and purpose of the research and asked for their consent to be subjected to semi-structured interviews before they and their staff were subjected to an on-line questionnaire. Following the semi-structured interviews, the professional staff of the auction houses were invited to be respondents in the on-line questionnaire. When the questionnaires were sent out, they were informed that consent to contact them had been sought from the heads of their respective auction houses. They were then invited to respond and were thanked in anticipation of their participation. The on-line questionnaire preserves the anonymity of the respondents. Once the research has been completed, all details of the respondents and their responses will be deleted.

5.9 Issues appertaining to tenure

Data relating to tenure became available during the cross-sectional study. These tenures were observed and recorded. The capital values between freehold and leasehold interests differ. All other things being equal, the norm is that leasehold interests should be valued by applying different yields from those applied to freehold interests (Isaac and Steley, 2000: 25-30 and 126). This is shown to be a reason for treating tenure as a variable having an influence on yield. Such an approach is confirmed by the findings of the qualitative study. The sponsors of this research are only interested in identifying the factors influencing yields with respect to freehold premises. Therefore, leasehold interests are excluded from the study.

Scotland is a geographical area that has a distinctly different legal system from the remainder of the United Kingdom. As a consequence, it has a different system relating to land law and tenures. It needs to be established which Scottish tenures should be treated as being akin to freehold if Scotland is to be included in the different research methods. This, therefore, needs to be considered at the research design stage. Hence, Scottish land tenures are considered in 5.10.

5.10 Scottish land tenure

In order to ensure internal validity, it is important that the study only includes those Scottish land tenures akin to freehold.

Scots law has always been very different to that in the rest of the United Kingdom. Whereas that in England, Wales and Northern Ireland is based upon the common law, that in Scotland has been based upon Roman Law. Furthermore, right until the turn of the last millennium, Scots land law was based upon feudalism. Feudalism had been its foundation for the preceding thousand years.

According to Paisley (2000), prior to the abolition of the feudal system:

> "All title derived from the Crown which retained the title as paramount superior. Under the Crown existed mid-superiors each of whom held a right known as *dominium directum*, and lastly there was a party (known as a "vassal") who owned the property right known as *dominium utile*. Only the vassal was entitled to occupy and use the land although the right of superiority conferred on the superior the most important of which in recent years was to enforce feudal real conditions. With the abolition of the feudal system in Scotland superiorities of all types will be extinguished and their rights of enforcement of feudal title conditions will disappear subject to limited preservation by various means including reallotment to a neighbouring tenement."

The feudal system in Scotland was largely brought to an end by the enactment of the Abolition of the Feudal Tenure etc (Scotland) Act 2000. This piece of legislation by the Scottish Parliament brought about the greatest change in Scots land law for almost a millennium. The 28th November, 2004 was the appointed day under the legislation for the final demise of the old forms of tenure. After this date all *dominium utile* titles become absolute forms of ownership free of feudal obligations, the intermediate interests having been abolished. Hence, before the implementation of the reforms under the Abolition of the Feudal Tenure etc (Scotland) Act 2000, the ownership of real estate was essentially feudal, whereas now ownership is outright. Accordingly, the former interests of the superiors have now disappeared. Outright ownership is now vested in the former vassals and their successors in title.

The 2000 reform of Scots land law had been necessitated by developments elsewhere. The enactment of the Human Rights Act 1998 by the Westminster Parliament gave effect throughout the United Kingdom to the European Convention on Human Rights (ECHR). The Human Rights Act 1998 did not necessarily repeal existing primary legislation that contradicted the ECHR (Paisley, 2000). Whereas the courts did not have the authority to strike out legislation enacted by the Westminster Parliament, the Scottish Parliament could not enact legislation that was contrary to ECHR rights (Paisley). This consideration appears to have been perhaps the principal impetus behind the abolition of the Scots feudal system. However, any reform had to be framed in such a way so as not to impinge on the ECHR rights of the former superiors.

Notwithstanding the legal distinction, Scotland feuholds and the interests of vassals had been treated as akin to freehold interests elsewhere in the United Kingdom. Therefore, where the terms feuhold and feudal have appeared in auctioneers' catalogues, these have been treated for the purposes of this research as being identical to freehold. The question arises as how to treat the term heritable which now appears as a descriptive term for Scots tenure in the catalogues.

The current position is that Scots law divides property into corporeal and incorporeal property. It further distinguishes these divisions between heritable and moveable categories. The result is that Scots law recognizes four overall types of property. The Scots real estate tenures that may appear in the tenure variable of the datasets to be tested in this research will fall into either corporeal heritable property or incorporeal heritable property. Leaseholds fall into the incorporeal heritable category. Land and buildings fall into the corporeal heritable category.

Corporeal heritable property is the term that Scots law applies to the proprietorship or ownership of real estate. Heritable tenure is not shown to be akin to freehold tenure. For this reason, the analyses of the datasets in this research will omit corporeal heritable property from the analyses of freehold property. This affects five cases observed and recorded in the original, full dataset comprising all tenures.

The position relating to Crofting Tenure, which relates to crofts in the seven crofting counties (Edwards, 2005) has not been considered, because no banking-halls in Scotland were found to fall in that tenure. Udal, an old Norse tenure relating to The Shetlands and Orkney would not impact, since it relates primarily to foreshore and seabed rights.

5.11 Conclusions

A research design is formulated based upon the four levels identified by Trafford and Lesham as being important elements. An appropriate paradigmic position is considered. Whilst the paradigm in this research is essentially a positivistic one, it is tempered by the fact that the study is undertaken within the area of social research.

During Stage 1 of the Professional Doctorate, emergent theory evolved from the case and pilot studies. Such emergent theory was used to suggest hypotheses to be tested in the current research. However, a review of the relevant literature is used independently from the Stage 1 studies to produce the theoretical perspectives. These identify hypotheses to be tested in this research. The testing, both by qualitative and quantitative study, of the factors identified by the theoretical perspectives makes this research even more robust through triangulation.

The positivistic nature of the current research leads towards both survey research and experimental research as a form of enquiry. They are respectively qualitative and quantitative methods. The use of both allows for methodological triangulation. Further triangulation is achieved within the qualitative study by splitting that into two parts: a semi-structured interview followed by a questionnaire. The quantitative research is a correlational one undertaken through cross-sectional study.

The research design has been formulated to minimize threats to validity and reliability. Validity has been enhanced by using a raft of theoretical perspectives to underpin the qualitative study, which in turn triangulates to quantitative study. Additional validity is afforded the quantitative study by additional testing using post-study data. Measures are put in place to reduce threats to reliability. In particular, measures are put in place to ensure that the data are correctly recorded and not subsequently corrupted by coding issues.

Ethical issues are primarily confined to the qualitative study. These have been considered and addressed.

Issues relating to tenure have also been considered. The investigation into British banking-halls includes premises in Scotland. The law on tenures is different in Scotland from the law on tenures in England and Wales. Land tenure in Scotland comparable with freehold is identified so that the former may be included in the study. In this way, threats posed by tenure to the validity of the research are eliminated by the research design.

The research design of the current study is founded on a paradigm slanted towards positivism and it is triangulated across methodologies and within the qualitative study. The whole is underpinned by the theoretical perspectives. As a result, a robust study is conducted into the main factors maximizing the investment yields of retail bank premises.

Table 5.1 summarizes the research design and methodology.

Table 5.1 Summary of research design and methodology chapter

Hypothetical relationship	Theoretical basis	How the influencing factors are operationalized	The source of data for the qualitative study	The source of data for the quantitative study
The positive relationship between yield and the influencing factors.	Academic literature and property investment reference books used to develop theory at generic level. Modern Portfolio Theory applied to property to show that diversification is adopted due to the presence of influencing factors on performance (Fraser, 2004).	Statistical methods capable of analyzing the predictor variables as categorical data are used.	Semi-structured interviews followed by questionnaire of experts in the field.	Published data of banking-hall investments sold at auction between 1997 and 2006. Post-study data used for further validation draws on the same sources.

6 The qualitative study: identifying the predictor variables

6.1 Introduction

Holmes et al (2006) define in the simplest of terms the distinction between quantitative and qualitative study. In essence, they define quantitative study as comprising numerical data and qualitative study as comprising descriptive data. Flick (2002) states that having been used in psychology since the beginning of the twentieth century, qualitative study has become a commonly used approach in social research. Qualitative study embraces several theoretical methods which, however, all derive from subjective viewpoints (Flick, 2002).

Data of the British retail banking premises sold at auction on sale-and-leaseback terms during the study period are published and are in the public domain. These data provide the basis of the quantitative study using empirical material. They have been analyzed to show associations, causes and effects so that models that explain and predict *Yield* and *Yield Group* respectively can be built. From these models, it is possible to generalize. However, it has been necessary to conduct the research in such a way that the appropriate variables were analyzed during the quantitative study. Qualitative study enables those appropriate variables to be identified beforehand (Huberman and Miles (2002). Although the use of deduction from existing theory has identified factors likely to influence *Yield*, qualitative methods in the form of interview and questionnaire have confirmed theory in the context of banking-hall investments and have also enabled the research to become more robust.

6.2 Sources of the qualitative input

The sale-and-leaseback of British commercial investment property by auction during the period was limited to five specialist property auction-houses based in London. Of these five practices, three were large auction-houses conducting auctions of several hundred properties at a time, whereas the smaller two auction-houses conducted a lower volume of sales. One of the smaller practices ceased trading during 2008 due to insolvency. It was established that the second smaller practice did not auction banking-halls on sale-and-leaseback terms and that it offered next to no such investments for re-sale. All of the three larger practices auctioned volumes of retail banking premises both during the study period and subsequently. The banking-halls sold at these auctions were a mixture of those sold on new sale-and-leaseback terms by the banking companies and those re-sold by existing investors. Nearly all of the banking-halls were auctioned through the three largest auction-houses.

This qualitative study was undertaken on the basis of data collected from the professional staff and partners within all of the specialist auction-houses able and willing to take part. It was achieved in two parts, each being a form of survey in its own right. Firstly, an interview was undertaken with heads of auction-houses willing and able to respond. Then a questionnaire was completed by the partners and other professional staff in those auction-houses willing and able to respond.

6.3 Validity of the qualitative study

Surveys are a way in which descriptive data may be collected from professional practitioners (Fink, 2009). However, before a questionnaire is put together, thought needs to be given to how the questionnaire will be structured and in particular how the questions will be framed. It is important that the qualitative study is built on sound foundations and is therefore valid.

Maxwell (1992) states that during the long-running debate about the legitimacy of qualitative research, validity has been at the core of the argument. Hence, those favouring a quantitative and positivistic paradigm have long criticized qualitative research for a lack of measures to ensure validity. These opponents attack the absence of scientific measurement, the lack of strict hypothesis testing and the deficiency in tight controls to counter risks to validity (Maxwell, 1992). To counter these arguments put forward by Kirk and Miller (1986) from the perspective of the positivistic paradigm, it is argued that qualitative research is able to ensure validity through different procedures (Maxwell (1992).

6.4 The use of surveys

The use of surveys is widespread. They are used not only by researchers, but also others, including business managers, administrators and those influencing social, economic and political policy (Fink, 2009). Data collected using qualitative methods often lack order and structure, and may be cumbersome (Huberman and Miles, 2002). Accordingly, Huberman and Miles argue that the qualitative researcher needs to order such data in a way that it still retains the original information. Fink shows that it is important that qualitative researchers using surveys must establish a survey's objectives and, therefore, the precise questions that lead to meeting those objectives. She also states that it is necessary to determine the correct sampling and design. Hence, it is necessary to determine who the participants will be and how the survey will be conducted (Fink, 2009).

Fink (2009) shows that both the types of survey and the way in which they are administered vary. She states that all types of survey comprise questions and responses. However, there is a continuum. At one end of the continuum there is the completion of the self-administered questionnaire, and at the other end there is the interview (Fink, 2009). Modern means of communication allow surveys to be conducted from a distance. Thus, questionnaires no longer have to be conducted face-to face or even by post. They may now be completed on-line. Similarly, interviews need not be conducted face-to-face, since they may be undertaken by telephone or video-conferencing.

6.4.1 The sampling of the qualitative study

The question arises of who should be included in the qualitative study. A decision has to be made as to whether to include everyone as a respondent or to just rely upon the data from a sample of respondents. Constraints upon time and other resources may mean that it is not possible to survey the entire population at large (Fink, 2009). Sampling is a means by which researchers may collect data without having to collect data from the entire population. Therefore, it is important to avoid sampling errors (Fink, 2009).

Fink (2009) states that where a random sample is taken for a survey, it is said to be *objective*. Samples taken randomly using a computer program would meet this criteria (Fink, 2009). A convenience sample is different: it is not objective. Fink (2009, p. 52) states:

> "A convenience sample includes people who are available and willing to take the survey."

Thus, if the respondents within professional practice, who participate in surveys aimed at identifying the predicator variables affecting the investment yields of British banking-halls sold at auction, are only those willing and able to participate, those surveys will not be considered as being objective.

The sale-and-leaseback of real estate assets is a very specialist area of professional practice. Moreover, a very limited number of real estate professionals work in the area of sale-and-leaseback of British banking-halls through the medium of auction. Therefore, a survey of these individuals based upon a random sample would be inappropriate. Instead, it is appropriate to base samples upon panels comprising experts (Fink, 2009). These types of survey respondents fall into what Fink (2009) describes as *convenience samples*. Such convenience samples may comprise professional experts who are capable of giving their expert views where the data are thus far unknown. Panels are a type of convenience survey that allows experts to be surveyed by interview or questionnaire.

When looking at convenience sampling, Bryman (2004) cites the example of sending questionnaires to teachers taking a part-time master's degree in education when researching the positive attributes of headmasters. Bryman states that this group of respondents has been selected simply due to its accessibility. However, he argues that the findings of such research could not be generalized, since it is not known of what population this sample of teachers is supposed to be representative. Nevertheless, Bryman does not entirely dismiss the effectiveness of convenience surveys. He states that convenience surveys are commonly used in social research. Moreover, he makes the case that convenience surveys are important in identifying predictors that may be capable of influencing the dependent variable. Bryman shows how convenience surveys can be used to formulate a raft of questions intended to measure teachers' preferred leadership attributes. Similarly, convenience surveys, based upon samples comprising specialist professionals, having expertise in the sale of British retail bank premises by auction, provide an effective way of identifying the predictor variables that influence *Yield*.

Since the sale-and-leaseback of British banking-halls by auction is such a specialist area, the sample used in the qualitative study is a convenience sample comprising the same number of professional experts, who are able and willing to be respondents. Convenience samples are able to draw samples from the wider professional community that only includes experts in the particular field. Thus others, who might be selected from within professional practice by one of the random sampling techniques, but who may well not be expert in sale-and-leaseback, are omitted. In adopting a panel of experts based upon convenience sampling, the validity of the qualitative study is ensured.

6.4.2 The design of the qualitative study

Survey design is important, because such design affects the validity of a survey (Bryman, 2004); Fink, 2009). According to Bryman (2004, p. 27), the research design may be encapsulated as:

> "A framework for the collection and analysis of data. A choice of research design reflects decisions about the priority being given to a range of dimensions of the research process."

An inappropriate framework for the collection and the analysis of the data will result in a loss in the validity of the survey findings (Fink, 2009). There are different types of validity. Both Bryman (2004) and Fink (2009) identify these as *measurement validity*, *internal validity* and *external validity*. As a social scientist, Bryman also identifies *ecological validity*.

Measurement validity relies upon the selection of the appropriate survey instruments (Fink, 2009). That is to say that measurement validity considers what is being recorded and whether such recording is relevant to the effect (Fink, 2009). Thus, in the case of the present research, it has to be considered whether the yield on a particular banking-hall is likely to be high or low. Bryman makes the point that measurement validity is associated with reliability. Therefore, in order to have measurement validity, the research must produce a model that is capable of explaining yield with a high degree of reliability. This research seeks to extend such reliability to the prediction of yield subsequent to the study period.

Internal validity is primarily concerned with causality. This amounts to more than a mere relationship between variables. Rather, it amounts to whether one variable actually causes an effect on another variable (Bryman). So, for example, where it could be shown that an increase in mortgage finance costs had caused a fall in capital values of properties subject to a study, that study can be said to be internally valid. The essence of validity is what is the level of confidence that an effect in the dependent variable has in at least been part attributable to a given predictor variable (Bryman)?

Whereas internal validity is concerned with causality within a given study, external validity is concerned with whether the results of such a study can be generalized beyond the study (Bryman). Hence, if the qualitative study was not externally valid, it would not be capable of being generalized beyond the sphere of the respondents. However, not only are all of the respondents property professionals, but moreover they are experts in the disposal of properties by sale-and-leaseback by auction. Therefore, it is entirely reasonable to expect the results from the respondents' participation in the surveys to be capable of being generalized. Furthermore, the results of the surveys forming the qualitative study are confirmed and triangulated by the subsequent quantitative study.

Bryman raises the issue of ecological validity. He argues that the more that the social scientist becomes involved in an experiment, the more the results may become ecologically invalid. This is because through the process of the investigation, the researcher might influence the outcome, since the respondents are not being studied in their normal environment (Bryman). However, the respondents in this qualitative study were surveyed in a context not dissimilar from their normal professional environment. Moreover, the respondents were surveyed in a field in which they were the experts and in which they should have felt entirely comfortable.

6.4.3 Survey methods adopted

Due to geographical dispersion, interviewing of survey respondents can be expensive and time consuming. Also, busy professionals do not always favour having researchers visiting their offices. Accordingly, postal questionnaires can provide an effective alternative means of conducting surveys (Bryman). On-line versions of questionnaires are now capable of being created and administered (Fink, 2009). On-line surveys generally generate a better response than postal sources (Bryman). However, Fink (2009) shows that it is important to pilot test and refine questionnaires especially to ensure validity. This can be achieved by trialling the questionnaire with individuals other than the respondents in the sample (Fink, 2009). Likely questions may be identified by telephone interviews with the leading experts.

The qualitative study in this research relied upon two surveys: the interview of the senior partner of each auction-house involved in the sale-and-leaseback of British banking-halls, followed by the sending of the final questionnaires to all the partners and professional staff working in those auction-houses.

During the study period, five property auction-houses based in London had undertaken the sale of banking-halls as investment properties by auction. Three of those auction-houses had dominated the market. The remaining two auction-houses had re-sold a small number of banking-halls previously sold by the others on sale-and-leaseback terms. One of the two smaller auction-houses ceased trading in 2008 and was, therefore, unable to participate in the surveys. The second of the two smaller auction-houses advised that it was unable to participate in the surveys due to having had such little turnover of banking-halls. The senior auctioneers of the three largest auction-houses were interviewed by telephone. The responses of these three were used to confirm and refine the questions within the questionnaire sent to the partners and the professional staff of those auction-houses surveyed. However, the senior auctioneer of one of these auction-houses indicated that he wished to complete one questionnaire on behalf of all of the staff within his department. That offer was rejected on the basis that the inclusion of data on that basis would have prejudiced the research and brought the validity of the qualitative study into question. Therefore, all the respondents were drawn from the partners and the professional staff of the other two main auction-houses.

Each of the respective heads of the three main auction houses participated in the semi-structured interviews. Therefore, the semi-structured interviews comprised three respondents, who represented one hundred per cent of the population of the heads of the main auction houses selling banking-hall investments.

The total population of expert professionals from within the three main auction-houses was nineteen. All twelve expert professionals from the two auction-houses that participated in the questionnaire responded. Therefore those who responded represented just over sixty-three per cent of the population.

Although the qualitative study was based upon a questionnaire completed by a sample comprising those experts able and willing to participate, it was still necessary to determine the precise questions and the way in which the questionnaire was to be administered. Huberman and Miles argue that triangulating by several methods of data collection methods makes theory and hypothesis more robust. Combining both quantitative and qualitative methodologies makes the research more robust (Flick (2002), Huberman and Miles (2002)). Huberman and Miles also go further than this inasmuch as they make a case for also using different methods of qualitative study in combination as an additional form of triangulation. They cite Bourgeois and Eisenhardt (1988) in combining interview with questionnaire as a way of confirming emergent theory. Accordingly, the qualitative study in itself is made more robust by combining methods, including interview with questionnaire.

The *self-administered questionnaire* is often called the *self-completion questionnaire* (Bryman). It has many similar attributes as the structured interview, except that it does not have an interviewer to put the questions. Therefore, the questionnaire needs to be easy to follow and complete (Fink, 2009; and Bryman, 2004). As already stated, the self-completion questionnaire is cheap and easy to administer, and, once refined, is capable of being completed by expert professionals without further assistance.

6.5 The conduct of the qualitative study

The qualitative study comprised two surveys. Together, these two surveys had three essential elements, which can be summarized as follows:

- The creation of a draft questionnaire based upon theory derived from the literature.
- The interviewing of the heads of auction-houses in order to ascertain their views on likely influencing factors.

- A self-completion questionnaire completed by experts in the field.

A draft questionnaire was compiled before being placed in person by the researcher before a small sample of chartered surveyors competent in the valuation of commercial investment property in the United Kingdom. These participants were competent property investment professionals, but not experts in the field of sale-and-leaseback. Initially, these participants were selected by the researcher. Other respondents were identified by the first participants. This is a technique known as *snowball sampling* and is a form of convenience sampling, and as such is another form of non-probability sampling (Bryman). The respondents from the snowball sample were in the main used to confirm the theory derived from the literature. They were also asked to add any further comments on what they might have considered to be relevant factors.

Once any relevant comments from the participants of the snowball sample had been fed back into a revised questionnaire, the senior auctioneers of the respective auction-houses were contacted by telephone. As outlined in 6.4.3, there were three leading auction-houses, based in London, which had dealt sufficiently with the sale-and-leaseback of British banking-halls to have the necessary expertise in the field. The three heads of the respective auction-houses were contacted by telephone with two objectives in mind. The first objective was to conduct a telephone interview with each head of the respective auction-houses. The second objective was to seek permission from each head, in order to comply with ethical considerations, to ask their respective members of professional staff expert in the field to participate in the questionnaire.

Of the three heads of the respective auction-houses, two were very helpful and co-operative, whereas the third was either unwilling or unable to participate in full. The third head did not want the professional staff in his department to participate in the questionnaire. Instead, he proposed that he should complete just one questionnaire on behalf of all those in his department. This offer was rejected primarily on the basis that such an approach would not have taken into account any differences of opinion that might have been expected to exist amongst the professional staff within the department. To have included a single questionnaire completed by the third head on behalf of all his staff would have adversely effected validity of the study (Fink, 2009; and Bryman, 2004). In particular, to have included it would have affected both internal and external validity. Its inclusion might have jeopardized internal validity on the basis that to have done so would have reduced confidence that given independent variables were having a causality on *yield* (Bryman). Its inclusion could well have been considered to have affected external validity on the basis that it might have reduced ability of the generalization of the findings beyond the particular study (Bryman).

Holmes *et al.* (2006) define an interview as being a meeting convened between a researcher and a participant. In interview, the participant is invited to respond to the questions and the prompting put forward by the researcher. Interviews may be conducted face-to-face. However, it is possible for researchers to conduct interviews through media such as telephone and the internet in order to save time and money in travelling (Bryman). Due to constraints imposed by distance, the interviews used in this qualitative study were conducted by telephone.

Interview types range from structured to semi-structured to unstructured (Kane and O'Reilly-De Brun, 2001). Interviews are a useful tool employed in social research as a method of obtaining from a respondent a great deal of information (Bryman, 2004). However, thought had to be given as to the most appropriate type of interview to be used in this research. Therefore, the types of interview were reviewed in order to ascertain the most appropriate to be used. Notwithstanding the choice of interview type, each type of interview needs to hang on an *interview guide* (Bryman). Bryman (2004: 324) states that the *interview guide* is a somewhat vague term, but defines it as:

> "The brief list of memory prompts to be covered that is often employed in unstructured interviewing or to the somewhat more structured list of issues to be addressed or questions to be asked in semi-structured interviewing."

Even unstructured interviewing relies upon an interview guide in the form of an aide mémoire (Bryman).

At one end of the interview-type continuum, the unstructured interview is an interview where the participant is given the latitude to comment in their own way about a particular topic (Holmes *et al.*, 2006). According to Bryman, it gives greater flexibility in the research process so that once some data start to emerge from the first respondents, the direction and the nature of the investigation can more readily be adapted. This contrasts with a structured survey, whether that be an interview or a questionnaire, which may be too rigid and incapable of adaptation. An unstructured interview is defined by Bryman (2004, p. 545) as:

> "An interview in which the interviewer typically only has a list of topics or issues, often called an *interview guide*, that are typically covered. The style of questioning is usually very informal. The phrasing and sequence of questions will vary from interview to interview."

At the other end of the interview-type continuum, the structured interview is an interview where each participant is asked the same series of questions as all the other respondents (Bryman, 2004; Holmes *et al.*, 2006) and the questions are of the type that may be used in a questionnaire (Kane and O'Reilly-De Brun, 2001). A structured interview is defined by Bryman (2004: 544) as:

> "A research interview in which all respondents are asked exactly the same questions in exactly the same order with the aid of a formal interview schedule."

Whereas the unstructured interview allows greater flexibility, the structured interview reduces error arising from the variation of the questions (Bryman). Another advantage of the structured interview is that it enables the answers to be collated for presentation and summary (Holmes *et al.*).

The semi-structured interview lies between structured and unstructured interviews. It does not follow a standard interview form. According to Kane and O'Reilly-De Brun, the semi-structured interview relies upon an agenda which covers the salient points and acts as an aide mémoire, but with the actual questions framed to the respondent and the situation. Furthermore, it is necessary to use professional language when interviewing the experts.The semi-structured interview permits the use of professional language and in-depth discussion with experts in the field. Kane and O'Reilly-De Brun argue that it is necessary to do so when interviewing the experts in order to build trust and confidence in their eyes. Kane and O'Reilly-De Brun also state that in conducting such interviews with the experts, the researcher should develop a brief list of a limited number of questions or build a mental map of what information is being sought. They further state that during semi-structured interviews, the researcher should be prepared to be flexible.

Bryman asserts that the term *semi-structured interview* covers a type of interview where the researcher draws the questions from the interview guide, but gives the respondent a great deal of freedom in responding. This means that the questions are typically more general than those used in a structured interview.

The different types of interview on the continuum of interview-types are used at different times in different elements of study. However, as Flick (2002) shows, semi-structured interviews are a good method of tapping into complex and expert knowledge on a topic. The present qualitative study has used interviews in order to create the questionnaire that was subsequently sent to respondents comprising the partners and other professional staff working in the London auction-houses dealing with the sale-and-leaseback of British banking-halls. Since the interviewees were the heads of those auction-houses and were the ones with the expert knowledge on the topic, it was appropriate to adopt a semi-structured interview. Therefore, the interviews of the heads of the auction-houses were conducted as semi-structured interviews.

6.6 The semi-structured interview and its responses

As highlighted in 6.4.3, of the three main London auction-houses dealing in the sale-and-leaseback of British banking-halls by auction, the heads of two of them indicated that they were willing and able to participate in the qualitative surveys. Since the head of the third auction-house had indicated that he wished to complete a single questionnaire on behalf of all his staff, that offer was rejected on the grounds that it could prejudice the validity of the qualitative study. Notwithstanding that, this consideration did not preclude the head of the third auction-house from being a participant in the interview survey. On the contrary, his expert views on the effects likely to have a relationship with the yields of banking-halls were just as sought-after as the views of the heads of the other two auction-houses. All three had expert knowledge that could be used to confirm the questions to be included in the final version of the subsequent questionnaire. Accordingly, the head of the third auction-house was included as a respondent in the interviews based upon the interview guide.

A semi-structured interview technique was adopted, because it is one that allows an in-depth discussion of the topic to develop with the expert respondents (Flick, 2002). The interviews commenced with each of the respondents being told that the researcher was a chartered valuation surveyor, that the purpose of the research was primarily to establish the effects apparently having an influence upon the yields of British banking-halls sold at auction and that they were being interviewed as the leading experts in the field. The respondents were then allowed to formulate their own responses with prompting, where necessary, based upon the questions in the interview guide shown in Appendix I. The questions in the interview guide were based primarily on the theory generated by the literature review. The respondents were asked to try to frame their responses in the context of the study period which was from 1997 to 2006.

The three leading experts, who were the respondents to the interview survey, concurred in their responses. Their collective responses are summarized as follows:

- Most of the purchasers of British banking-halls sold at auction on sale-and-leaseback terms were private investors.
- The main influences having an effect upon *Yield* were:
- Region.
- Lot size.
- Being one of the main retail bank brands rather than a secondary retail bank.
- Investors were looking for 'good, solid, blue-chip covenants', hence the preference for the main retail banks.
- Investors sought longer lease reversions; fifteen and twenty year terms were especially popular.
- Investors favoured main town centres and prosperous market towns as opposed to secondary areas and suburbs as locations.
- 'Good, solid buildings' were favoured by investors.

During the interviews, the respondents were encouraged to elaborate, with slight prompting, on the underlying reasons influencing investors' choices. The collective reasons given by the respondents can be summarized as follows:

- Not only banking-halls, but many classes of sale-and-leaseback properties have been popular with investors in the United Kingdom, because those properties have been viewed as a safe and secure medium of investment. This has been especially so with low rates of interest on deposited capital, volatility in the stock markets and failing pension schemes in recent years. Readily available mortgage finance at relatively low rates of interest prior to the start of the global financial crisis in 2007 helped to facilitate greater participation in sale-and-leaseback investments. As a class of sale-and-leaseback investment, British banking-halls have been an especially safe and secure form of investment.

- Economic factors peculiar to a given region would have some impact upon the overall yields of that region. However, within any given region, there would be discrepancies between locations due to the very local social and economic factors of those locations. Such precise observations accord with the literature. Furthermore, most private investors preferred to purchase property investments within reasonable proximity to their home addresses and therefore within their own regions. As a result, clusters of private investors in a given region tended to help to drive yields for banking-hall investments down in that region.
- Generally, there was a lot-size effect. Although some wealthy individuals were not constrained by lot size in terms of price, most investors were. Such constraints were due to a combination of ready cash and the availability of mortgage finance to respective investors. Hence, although most investors preferred to purchase better and, therefore, more expensive premises, an upper price limit was generally imposed by the availability of finance.
- Investors had a preference for the principal retail banks rather than the secondary banks, because they were seeking a safe and secure haven for investment. Those investors investing in banking-halls were in particular seeking reduced risk and greater security.
- The reason why main town centres and prosperous market town locations were favoured was that such locations were considered more likely to retain their banking-halls over time. Those rationalizations of banking-halls that had taken place primarily occurred in the peripheral and less prosperous locations.
- Many private investors had an irrational desire simply to own a solid building that had the traditional attributes of a banking-hall.

Perhaps surprisingly, the three respondent experts made no reference to the adaptability of the buildings to uses other than retail banking. Bearing in mind the announcement by HSBC that it was to trial new, glass-fronted retail banking premises which appeared to be more user-friendly and in which staff could mingle with customers (Chesters (2006), Poulter (2006) and HSBC Holdings plc (2006)), the respondent experts made no mention of such premises being favoured by investors.

6.7 The questionnaire and its responses

For ethical reasons and so as to retain their goodwill, the respondent experts who participated in the interview survey were asked for their consent for all the professional members of their staff, including themselves, to be sent a self-administered questionnaire. The heads of two of the auction-houses readily agreed to give their consent. Since the head of the third auction-house wanted to complete a single questionnaire on behalf of all his professional staff, his offer was declined. Accordingly, self-administered questionnaires were sent to all the partners and other professional staff from within two of the three main London auction-houses selling British banking-halls on sale-and-leaseback terms.

The questionnaire sent to the respondents was refined following the qualitative interviews with the heads of the three main auction-houses. It is reproduced in Appendix II. The questionnaire was created using an on-line software package known as *Survey Monkey*. The Survey Monkey on-line software package allows surveyors to use templates to create questionnaires in digital format, which can then either be conducted electronically from a computer or printed-out for manual completion.

Bryman states that on-line surveys are an effective way of conducting social research, because they are a cheap and quick way of surveying many respondents without constraints imposed by distance. He lists the shortcomings of this approach as being the lack of universal access to the internet, e-mailed questionnaires being ignored as an unsolicited intrusion and the inability to experience a respondent's tone and body language. However, these concerns are not applicable to the present questionnaire. All professional staff at the two participating auction-houses were furnished with questionnaires, having been advised that the heads of the respective auction-houses had given consent for their participation. Since, the respondents were experts being asked to complete a questionnaire about their field of expertise, the issue of tone and body language did not arise.

The researcher only had available the e-mail addresses of all the professional staff within one of the two participating auction-houses. Accordingly, all the respondents within this auction-house were sent an e-mail asking them to use an internet link through which they could respond to the questionnaire electronically on-line. These are normally known as web surveys (Bryman). All those invited respondents responded using the internet link. In the case of the second participating auction-house, printed copies of the questionnaire were posted to the head of that auction-house. The head of that auction-house then distributed these versions of the questionnaire to each of his professional staff. These were quickly completed and returned to the researcher in order to be manually entered into the Survey Monkey software package.

The questionnaire was promptly completed and returned. It appears to have had either a hundred per cent response rate or one very close to it from the respondent experts drawn from the two participating auction-houses. The results of the responses to the questionnaires were generated using the Survey Monkey software package. They are reproduced in the form of both as a case summary in Table 6.1. and as a bar-chart in Figure 6.1

Table 6.1 Percentage per factor identified by respondents as most likely to influence the yields of retail bank premises sold at auction on sale-and-leaseback terms

Case Summaries

	Percentage of responses	Selected Variables
1	100.0	Regional Location
2	91.7	Lot Size (as in hammer price)
3	91.7	Tenant bank company/brand
4	83.3	Tenure
5	16.7	Other (Please specify)
6	8.3	Physical size of the building
7	8.3	Proximity of rival bank groups
8	.0	Age of the building
9	.0	Brand cross-over within a group
10	.0	Close proximity of branches of the same brand

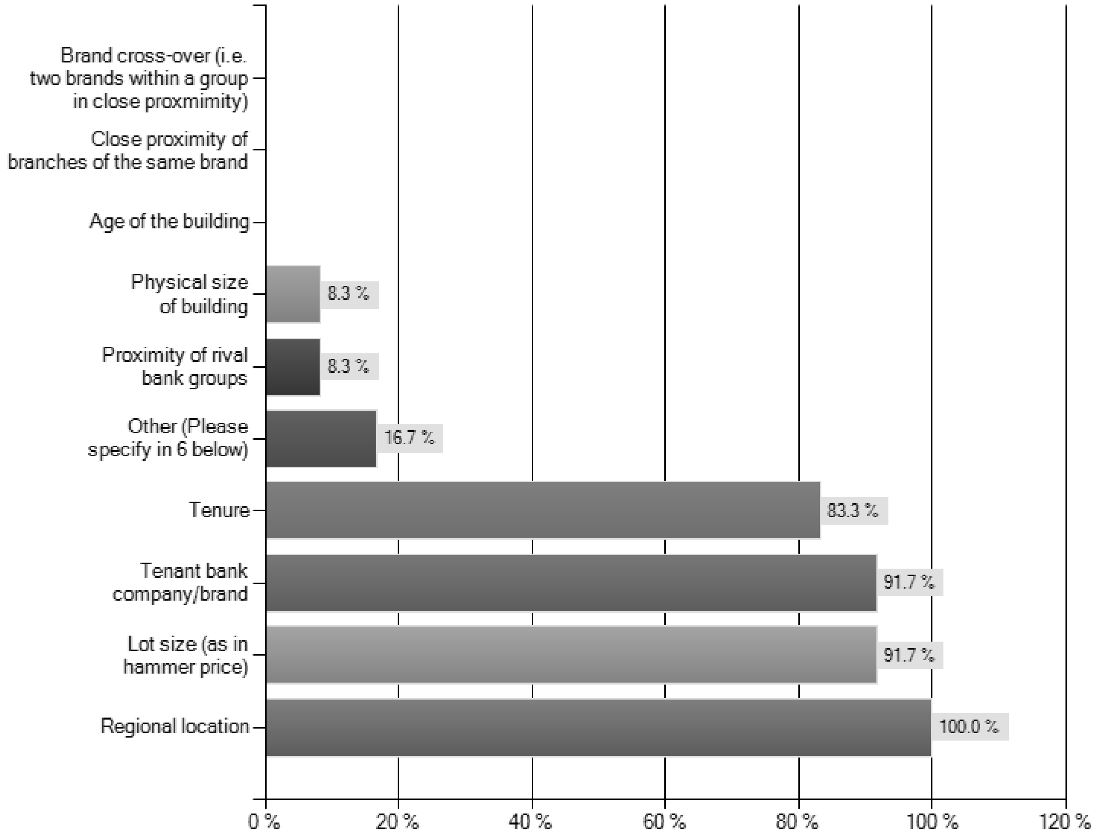


Figure 6.1 Percentage per factor identified by respondents as most likely to influence the yields of retail bank premises sold at auction on sale-and-leaseback terms

Four factors were identified by the respondents as having the greatest influence on retail bank premises sold at auction on sale-and-leaseback terms. These four factors all scored between 83.3 per cent and 100 per cent respectively. They were the same four factors that had been identified in the theoretical underpinning as having the greatest influence on yield. Only three other factors scored at all. *Physical size of the building* and *proximity of rival bank groups* only scored 8.3 per cent, which meant that these factors were only selected by one respondent respectively. Each of these two latter factors was selected by different respondents. The only other variable that was selected was *Other*.

Respondents who selected *Other* were asked to specify in the section of the questionnaire provided for all respondents to add comments about variables, other than those in Table 6.1 and Figure 6.1, which they considered as having a significant bearing on investors' investments decisions with respect to the banking-halls subject to the study. Six respondents, half of the total number, took the opportunity of listing any other significant effects. Most of these, making a sizeable minority of the total number of respondents, cited lease length. However, from these responses no clear lease length was discernible. The only clear pattern was that longer leases were more sought after than shorter ones by investors. Two respondents said that a lack of break-clauses was significant. No other effects were identified by any of the respondents.

Respondents were further asked for any additional comments or observations. None was made.

6.8 Discussion

Participation in the qualitative study was based upon a convenience survey of the experts in the field. Only twelve respondents participated in the qualitative study. The heads of the two auction-houses that participated were interviewed. These two heads were subsequently joined by the remaining partners and professional staff within their departments in responding to the questionnaire. All of the partners and professional staff from within the two participating auction-houses appear to have responded. Thus all those respondents willing and able to respond appear to have done so. Response to the survey was limited to twelve experts in the field. However, the validity of the qualitative study has been ensured by limiting it to a panel of experts using convenience sampling. Furthermore, those experts willing and able to respond constituted over sixty-three per cent of the entire population of experts.

The results of the qualitative study show that the four factors most likely to have influenced the yields and investors' decision-making when purchasing retail banking premises at auction on sale-and-leaseback terms between 1997 and 2006 were:

- Tenant bank company and brand
- Regional location
- Lot size as defined by hammer price
- Tenure

The four factors selected as having been likely to have been most influential received a response of 83.3 per cent or above, whereas no alternative factor received a response rate exceeding 16.7 per cent. Only two alternatives scored at all. The higher of these two alternatives was the *Other* variable. All respondents, including the two respondents who selected *Other*, were permitted to add further comments about additional significant factors. The main theme to emerge from these comments was that longer leases were favoured by investors over shorter leases. However, respondents were not specific about the optimal lease length.

Since the latter part of 2007, there has been a global financial crisis known as the *Credit Crunch*. Results of the sales of banking-hall investments sold during the Credit Crunch shows that even during that period, yields for this class of investment remain low relative to alternatives (Allsop, 2009).

In spite of the turmoil caused by the Credit Crunch, the demand for investment property in the United Kingdom post 2007 has remained high (Walker, 2009). Walker states that this high demand has been due to low rates of interest on capital deposited, resulting in property yielding between five and ten times the return of deposits. As a result, there has been renewed market activity by private investors during the credit crunch especially for property investments with the strongest covenants (Walker). Since bank tenants have always been perceived by the market as being strong covenants, much of the revival in demand from private investors has been directed towards banking-halls.

Reasons highlighted in the literature as being influential on the factors were confirmed by the expert interviewees. However, the interviewees did suggest further, possible influences not highlighted in the literature.

Quantitative analysis becomes much more difficult as the number of dependent variables increases. The use of many variables often results in sparse tables, which can inhibit analysis (Agresti, 1996). For this reason, the four factors shown by the qualitative study as being most influential should be selected for testing in the qualitative study. The emerging theme about the possible effects of lease length could be the subject of future research.

Only 16.7 per cent of respondents stated that lease length was one of the four factors most likely to have influenced yield. In that respect, lease length is not a significant factor in terms of this research. However, a significant number of respondents did say that lease length would be a further consideration beyond the four main factors. Nevertheless, those respondents who did cite lease length, including those two respondents who listed it as a main factor, were inconsistent in the precise length.

The original, full dataset of 1,012 cases only included 30 cases comprising leasehold premises and no missing data. The full dataset was very heavily skewed towards those cases comprising freehold properties. Since the sponsors of this research had stated their interest in only freehold banking-halls, the small number of cases comprising leasehold premises is excluded from the quantitative study.

6.9 Conclusion

The results of the qualitative study confirm the four factors identified in the theoretical perspectives as having the greatest influence on the dependent variable *Yield*. Since the sponsors of this research wish to identify how the yields might be maximized in the yields of freehold banking-halls when building investment portfolios, *Tenure* was removed from the quantitative study. *Tenure* was removed from the quantitative study as being superfluous in terms of the remit of this research. Its removal also facilitates better analysis of the quantitative data.

Accordingly, the three predictor variables identified by the qualitative study for use in the quantitative analysis are based upon:

- Tenant bank company and brand
- Regional location
- Lot size as defined by hammer price

7. Introduction to the Statistical Analysis

7.1. Introduction

The data comprise details of the sale-and-leaseback of British retail bank premises between 1997 and 2006. There are no published data of banking-halls having been sold at auction on sale-and-leaseback terms in Northern Ireland during the period of investigation. The observed, collected and recorded auction data included, amongst other things, the rents reserved on the banking-halls and the hammer prices at which they are sold. This research seeks to identify approaches that property investors might adopt in order to maximize their yields (the ratio of rents to hammer prices) from investments in freehold British banking-halls.

7.2 Data collection

The data were collected in three ways:

- From on-line data sources
- From published journals
- By the researcher attending auctions

Since the beginning of the current millennium, extensive data on commercial premises sold at auction in the United Kingdom have been published on the internet. These published data include the results of banking-halls sold on sale-and-leaseback terms in England, Scotland and Wales. From time to time, some of the auctioneers have published some of the data on their requisite websites. Other on-line sources have provided data for those who have paid to subscribe. The market and professional advisors have used these sources of data with confidence. The individual auction houses have appreciated that the data recorded on these websites must be accurate for the market to retain confidence in them and for them to avoid potentially expensive litigation. As a result, this research has relied on the data provided by these sources with confidence. However, all of the financial data relating to British banking-halls sold at auction on sale-and-leaseback terms through the principal auctioneers from 2000 onwards has been published on-line, free of charge, at www.propertyauctions.com. The website does not reveal the identity of either its owner or its operator. However, a search was made on the internet via *Whois* in order to ascertain the identity of its registrant. This website was registered to Opennet (UK) Limited, which was listed at Companies House as a hardware consultancy and as having changed its name in December, 2003 from Information Systems Support Limited. For many years, the website has been used by valuation members of the surveying profession, who have considered it a useful and reliable tool. The website provided access to a copy of the original auction catalogue entry of the lot concerned. Each entry included the rent to be reserved for each property on completion of the sale-and-leaseback transaction. Since the website, in most cases, gave the sale price, often referred to as the *hammer price*, the yield could easily be calculated.

For auction data prior to 2000, reliance needed to be placed on archived printed format. Empirical data on the disposal at auction of British banking-halls on sale-and-leaseback terms existed in the public domain. Many property auction results were published in the journal *Estates Gazette*. Those results were not exhaustive, but sometimes permitted cross-checking where some of the on-line data were missing. For example, three tranches of banking-halls occupied by Yorkshire Bank were sold on sale-and-leaseback terms at auction during 2000. For some reason, the data relating to these lots were blanked-out on the on-line resource. However, the empirical data relating to two out of these three tranches were published in *Estates Gazette* immediately after those properties were sold. The rents reserved and the hammer prices for each of the lots sold in these three tranches were left blank on the on-line sources. Nor were the yields for these lots disclosed on the on-line sources. However, the on-line sources did reveal the addresses of the premises and the dates on which they were sold. Since the journal *Estates Gazette* published the results of some properties sold at public auction in the United Kingdom, a search was made of that journal. A search of the auction results pages of *Estates Gazette* did reveal the rents and hammer prices of the Yorkshire Bank lots sold at two of the three auctions. Therefore, the cases from those two auctions for which the data were recoverable were included in the *Purified Freehold* dataset. However, the cases appertaining to the Yorkshire Bank premises sold on 23rd May, 2000 were omitted from the *Purified Freehold* dataset on the grounds that the relevant data could not be recovered.

The researcher did attend, observe and record data into catalogues during the latter part of the 1990s. Auction data on a number of banking-halls sold during that period were observed and recorded by the researcher. Unfortunately, most of the archives containing that data were destroyed before the start of this research. However, a catalogue of ninety-nine Midland Bank premises, since re-branded as HSBC, sold at auction in London in 1997 has been recovered from the archive. Ninety-seven of the banking premises in that tranche were freehold. The data with respect to those ninety-seven premises was observed and recorded by the researcher at the auction. Accordingly, these cases were entered in a Field Book and included in the full dataset used in this study.

On observation, the raw, empirical data was entered onto data collection sheets. After obvious errors had been corrected, these data were copied into the Field Book. Most of the errors related to a small number of properties having been placed in the wrong region when recorded by the auctioneers and their agents. Most of these geographical errors related to a few premises in Devon and Cornwall having been erroneously recorded as being in Wales. From the Field Book, each case was entered into a dataset entitled, *Original Raw Dataset*.

7.3 Description of the data

The data collected and recorded in the Field Book was a mixture of numerical and categorical data. The variable *Price* was the hammer-price, that is to say the sale price at which the auctioneer knocked down with the gavel the property as being sold. *Price* was a continuous, numerical variable. Similarly, *Rent* was a continuous, numerical variable. These two variables were not used directly in their existing form in the analysis and model-building. Nonetheless, they remained important components. The variable *Yield* was the ratio of *Price* and *Rent*. It is a numerical variable and was arrived at by dividing the *Rent* by the *Price* and then multiplying the answer by 100.

According to Mackmin (2009), traditional valuation practice in the United Kingdom has reflected yield as the reciprocal of years' purchase. The concept of years' purchase has existed in valuation practice in England since at least the middle of the sixteenth century when the Court of Augmentations was charged with the task of selling the lands of the dissolved monasteries on behalf of the Crown (Scorgie, 1996; Tipping, 2006). Years' purchase is the number of years multiplier applied to the rent to calculate the capital value (Mackmin). Mackmin states that the reciprocal of this produces the yield normally adopted by the valuation profession in the United Kingdom. He states that this type of yield can variously be called the initial yield, the all risks yield or the capitalization rate.

Isaac and Steley (2000) state that with respect to freehold property let at the full market rent, the All Risks Yield (ARY) is the same as the initial yield. They describe the full market rent as the rack rent. They also show that the initial yield (or ARY) for freehold property is:

$$\text{ARY} = \frac{\text{rack rent}}{\text{purchase price}} \text{ x } 100\%$$

Where the leading London-based property auctioneers have published yields for lots sold, they have done so as the reciprocal of years' purchase and have normally referred to them as initial yields. All the yields calculated for use in this research have been calculated on the basis of being initial yields. Therefore, this research with respect to freehold banking-halls has relied upon *Yield* as being defined as the initial yield as shown by Isaac and Steley.

Other variables such as *Banking Company*, *Region* and *Lot Size* are categorical variables. Date is regarded as a numeric variable, but it and the derivatives *Quarter* and *Year* may be treated as either numeric or categorical. When analyses and model-building were undertaken using such categorical data, appropriate statistical methods needed to be adopted in order to facilitate useful and meaningful results (Agresti, 1996; Field, 2005).

The datasets comprised many variables. Property Investors do not consider single factors in isolation when evaluating potential investments. On the contrary, rational investors consider all the available factors together (Isaac and Steley, 2000). The auction data for British banking-halls sold as property investments comprise several variables that may have an influence on the *Yield*. The qualitative study (Chapter 6), underpinned by theoretical perspectives in Chapter 5, confirmed the four main predictor variables originally identified through professional practice. These were: *Region*; *Lot Size*; *Banking Company* and *Year*. The variable, *Tenure*, was omitted from the study, because this study is concerned with only freehold British banking-halls. Further studies might consider investigating how the yields of leasehold premises might be predicted. However, it needs to be borne in mind that for any quantitative study into leasehold premises to be undertaken, it would be necessary to have sufficient cases, and therefore sufficient data, to adequately fill the cells in a contingency table.

For the purposes of this study, it has been assumed that the English and Welsh banking-halls subject to the analyses are freehold as defined by being in Fee Simple Absolute in Possession. The auctioneers' description of the properties being freehold was assumed to mean this definition. For the reasons outlined in 5.8 and 5.9, feuhold premises in Scotland have been grouped with freehold properties in England and Wales for the purposes of this study. Any premises where part was let or sub-let for non-banking purposes were excluded from the study on the basis of being incapable of quantifiable comparison.

Region was defined as being the same geographical sub-division of Great Britain as used by the principal, London property auctioneers. Accordingly, this research placed each case in the same region as identified either in the respective auctioneers' catalogues or on-line by www.propertyauctions.com. The exception to this rule was where a very small number of cases had been placed in the wrong region when recorded in the primary data. In these cases, they were placed in the correct regions before being recorded on the data collection sheets and entered into the Field Book. Since April, 1997, the Standard Statistical Regions have been replaced by the Government Office Regions (GORs) as a way of classifying regions in England (Parnham & Rispin, 2005). It was useful in statistical analyses to follow these regions. However, the property auction houses based in London had their own classifications. The main difference between the GOR and auction house classifications was that the auction houses grouped together the GOR regions of *North-East* and *Yorkshire and Humberside* into one region labelled *North-East*. For the sake of consistency, the analyses in this research defined the regions as having the same geographical boundaries as those applied by the auction houses. These regions were as follows:

- East Anglia
- East Midlands
- London-M25
- North-East England
- North-West England

- Scotland
- South-East England
- South-West England
- Wales
- West Midlands

The region labelled by the auction houses as *East Anglia* closely corresponded with the region labelled in the GOR classification as *East of England*. The map of the GOR classifications for the period subject to the study and produced by the Office for National Statistics (ONS) is reproduced in Figure 4.1.

Lot Size has been a term frequently used by professional advisors and investors alike. It was also identified by theoretical perspectives and the qualitative study as having a significant effect. The term *Lot Size* has no direct relationship to the size of the property. Rather, it is used by property professionals and investors to categorize the value of the property. To the layman, the term *Lot Value* might be easier to comprehend. However, this research retained the term *Lot* Size on the grounds that to do so was in accordance with professional practice. Normally, it is categorized into three sizes: small; medium and large. Generally, such categorization is based upon capital values rather than rental values. The market does not have clearly defined parameters for each category of *Lot Size*. Also, the sizes of lots change over time. During the period subject to investigation, they became larger. However, at some other times, usually when the market has been in decline, property lot sizes have become smaller. For the purposes of this study, it was necessary not only to take into account the general approach to categorizing *Lot Size*, but also the range over the period so as to enable analysis. These considerations did produce sensible thresholds at which to cross from a lower lot size to the one immediately above it. Once these thresholds had been defined, it was ascertained from further research that each threshold was only one penny below the thresholds applied for the three bands of capital value for the levying of Stamp Duty Land Tax. This tax is levied *ad valorem* on purchasers on acquiring an estate in real property. The *ad valorem* amount increases as the capital price paid moves from a lower priced banding to a higher one. In case Stamp Duty Land Tax had an additional effect, the thresholds of the *Lot Size* categories were each moved by one penny so as to coincide with the tax thresholds. Accordingly, the three categories of *Lot Size* were divided as shown in Table 7.1. During the first half of the period subject to the study, Stamp Duty Land Tax was only payable on properties sold at over £60,000 (H.M. Revenue & Customs, 2005). However, no cases were sold at a price of £60,000 or below. From 1st December, 2003, with respect to non-domestic property, Stamp Duty Land Tax only became payable on properties sold for more than £150,000 (H.M. Revenue & Customs, 2010a). This remained so for the rest of the period subject to the study. Only one property was sold at below £150,000 during the latter part of the study period when the Stamp Duty Land Tax threshold was set at that level. That property was sold for £125,000 just two days after the date on which the tax threshold was raised. During the period subject to the study, the rates at which Stamp Duty Land Tax was levied were raised. However, the thresholds at which the tax was levied remained the same for the middle and upper price ranges.

Table 7.1: Categorization of *Lot Size*

Category	Hammer price lower limit (£)	Hammer price upper limit (£)
Small	0	250,000
Medium	250,000.01	500,000
Large	500,000.01	No upper limit

Banking Company was the label given to each respective retail banking brand occupying the respective premises. Some of the retail banking groups operating in the United Kingdom have done so through more than one brand. For example, Barclays Bank and Woolwich Bank are two banking brands of the banking group called Barclays Bank. Likewise, Lloyds TSB Bank group operated through the two brands, Lloyds TSB and Cheltenham & Gloucester. Subsequent to the study period, Lloyds TSB Bank group and Halifax Bank of Scotland were merged as Lloyds Banking Group due to the effects of the Credit Crunch. In each case, the respective brands have been treated separately and ascribed to different category with the variable *Banking Company*. Since HSBC formerly traded as Midland Bank, all HSBC and Midland Bank cases have been labelled as *HSBC* and placed within a single category. The banking companies for which data were observed and recorded were classified as:

- Alliance & Leicester
- Abbey

- Barclays
- Bradford & Bingley
- Bristol & West
- Cheltenham & Gloucester
- Clydesdale
- Halifax Bank of Scotland
- Halifax
- HSBC
- Lloyds TSB
- National Westminster
- Royal Bank of Scotland
- Woolwich
- Yorkshire Bank

Each case in the full dataset had the variable *Date*, showing the precise date on which the lot was sold. In this form, *Date* was a numeric variable. Other variables derived from *Date* were *Quarter* and *Year*. Each year may be divided into four quarters, each comprising three consecutive months. *Date* can also be transformed into another variable *Year*, in which all the lots sold in any given calendar year are combined. Making the transformation into *Year* could facilitate some analyses that might not otherwise be possible with *Quarter*. The reason was that the variable, *Year,* comprised fewer categories than *Quarter* did. Hence, each category within the variable *Year* would comprise more cases than the categories within the variable *Quarter* would.

Notwithstanding the different ways in which *Date* may be represented, the essential point is that all such versions represent time. Time is important, because over time the macro-economy and market risk are cyclical. The macro-economic cycle and market risk do result in a fluctuation of property investment yields over time (Dunse et al. (2007); Krystalogianni and Tsolacos (2004)). This remains true whether time is represented by *Date, Quarter, Year* or any other unit of time. Therefore, these measures of time are also a measure of economic performance. In the wider macro-economy, there are economic indicators that may be used where appropriate. However, in the property investment markets, the norm is to rely upon more specific property-related economic indicators to assess the performance of yield. The

Investment Property Databank (IPD) produces a number of property indices, including some that relate to United Kingdom property. In particular, the IPD has produced indices relating to the initial yields of United Kingdom retail property over time. Banking-halls are a sub-type of retail property.

7.4 Objectives.

The research aimed to identify a predictive framework by which investors in British banking-halls sold on sale-and-leaseback at auction might maximize their yields on such properties. Qualitative research can be used to identify trends in the operational property requirements of the retail banks. It can be used to identify the design, size and locations of banking-halls that the banks are likely to want to occupy in the future. It can also be used to identify the business plans and future deployment of resources within banks with a view as to how these influences might affect design, size and location of future premise needs. Such qualitative research needs to be further supported by the analysis of the empirical data derived from the auction results. Quantitative research enabled the empirical data from the auction results to be analyzed and used to create a predictive model. These analyses enabled models to be built and tested. Models that are both explanatory and predictive were sought. The objective was to have a validated model that could predict the yield arising from any given combination of categories from the predictor variables. With such a tool, investors and their professional advisors will be better placed to build portfolios of banking-hall investments on the basis of predicting which premises are likely to achieve higher yields.

7.5 Model-building Procedure

The auction data were examined in order to ascertain the most appropriate form of quantitative analysis. Since the recorded data had not arisen from a designed experiment, it was necessary to undertake an Exploratory Data Analysis to elucidate the structure of the dataset (for example, how many observations for each combination of predictor variables and the distribution of yield values) and to obtain indications of likely relationships between variables. Only after this could appropriate models be selected and fitted to the data. The general procedure adopted was as follows:

1. Exploratory Data Analysis
2. Data mining
3. Conclusions
4. Recommendations

The data mining component itself comprised several distinct phases which included:

1. Model Selection
2. Model fitting
3. Model testing
4. Inference

Data mining is now widely used by many large organizations in order to enable them to derive valuable business intelligence from extensive databases (Berson, 2000). It allows the analysis of datasets derived from the observation of uncontrolled real world events and not from designed experiments. According to Kao and Zhang (2003), with the availability of modern computers, researchers can apply data mining to areas of business that include financial forecasting. They show that due to the potential to extract meaningful and valuable information from large databases, large retailers, ranging from supermarkets to mail order companies, extract such information to profile customer groups for marketing and to reduce churn. In the financial sector, data mining is now widespread, where it is used for identifying the likelihood of credit card fraud and insider trading and for identifying insurance risk (Edelstein, 1997). According to Ranjan and Malik (2007), such data mining techniques also have useful applications in academia.

S.M. Weiss and N. Indurkhya (1998) identify investment analysis as an area suitable for data mining. Moreover, they state that it is an evolving area, which in particular can be used for prediction. In this context, it uses data from existing cases to predict the outcome of future cases.

8 Exploratory Data Analysis

8.1 Analyzing the data

Since the data within the main dataset were observed, collected and recorded from banking-halls actually auctioned within the context of professional practice, it was not possible to rely upon a designed experiment. Therefore, initial, exploratory data analysis was important in order to ascertain the structure and characteristics of the data and to ascertain if the existing data were suitable for model building. This exploratory data analysis was undertaken in the following stages:

1. A review of what was known of the origins of the data, including the listing of any idiosyncrasies

2. Review of summary statistics:

- Count
- Mean, variance, quantiles of *Yield* as the dependent variable
- Numbers at each level of the predictor variables of *Lot Size, Region, Bank* and *Year*
- Missing data

3. Cross-tabulations with *Yield* converted to a categorical variable

4. Lineplots of mean *Yield* against *Lot Size, Region, Bank* and *Year*

5. Boxplots of the dependent variable *Yield* against the predictor variables

6. Testing for normality and constance of variance using data from cells with large counts

7. Log-linear analysis of the contingency table with *Yield* as a categorical variable

Throughout the exploratory data analysis, conclusions were drawn from the results at each stage in order to enable these to be acted upon as appropriate in the succeeding work. For example, where there were missing data, consideration was given to the removal of the corresponding observations from the analysis. At the end of the exploratory data analysis phase, further conclusions were drawn and consideration was given to the implications that these would have on the subsequent phases of the analysis. In particular, it gave an indication of the types of model that it would be appropriate to investigate and which datasets should be used.

The full, raw dataset, which included all tenures, is reproduced in Appendix III. It provided a rich seam, which was capable of being mined so as to identify the best way in which a model could be built and validated with respect to the rental yields of British banking-halls. Such a model built from this empirical data is not only explanatory, but also predictive. It is explanatory, because it provides the best model to explain how the dependent variable *Yield* was influenced by the predictor variables. It is predictive because it also allows for the prediction of the yields of British banking-halls sold at auction after the study period. Such prediction includes the prediction of premises to be sold at auction in the future. A predictive model allows professional advisors and investors to anticipate what the rental yield on a freehold banking-hall is likely to be given a particular set of predictors. The prediction is based upon probabilities. Therefore, the model cannot predict with certainty what the yield of any one property will be. However, the predictive model remains a very useful tool in portfolio building, where the objective is to maximize rental yield.

8.2 Data Review

8.2.1 The full original dataset

The full original dataset was entitled the *Original Raw Dataset*, comprised 1,012 cases and included all tenures. Not only did it include freehold properties, but also leasehold and heritable tenure lots. Although the qualitative study did suggest that the variable *Tenure* had an effect on the *Yield*, *Tenure* was not identified as one of the four main predictors. It did not include properties which were only partly used as retail bank premises and which therefore were not true comparables. The full dataset also included cases comprising lots occupied by retail banks and building societies that form only a very minute number of the cases. Hence, even those banks and building societies comprising only one or two cases were included in the full dataset.

Most of the British banking-halls sold on sale-and-leaseback at auction during the period of investigation were sold by three of the five largest United Kingdom retail banks. Moreover, even the disposals by these three banking companies were unevenly distributed throughout the period. Disposals of premises by the other banks were also unevenly distributed throughout the period. Disposals by some of the smaller banking companies were concentrated in specific regions due to those companies tending to have regional, rather than national, coverage.

Table 8.1 shows how the *Original Raw Dataset* was sequentially collapsed during the Exploratory Data Analysis process.

Table 8.1 The sequential removal of cases to create the *Purified Freehold* Dataset

Name of dataset	No of cases in dataset	Remarks
Original Raw Dataset	1,012	All cases of premises let entirely as retail bank premises and including all tenures.
Price Included Dataset	932	The number of cases for which *Price* (realization) was recorded. 80 cases for which *Price* was missing are removed.
Low Counts Excluded Dataset	909	After a further 23 cases comprising banking companies scoring less than 5 counts were removed.
Purified Freehold Dataset	874	All freehold, and in Scotland feudal and feuhold premises, for which rents reserved and hammer prices were recorded (after the removal of 35 cases comprising 29 leasehold, 5 heritable and 1 unknown).
Three Main Banks Dataset	691	All freehold premises for which all data were recorded with respect to the three banks with the highest counts.

8.2.2 Purification – missing and anomalous data

Notwithstanding that investors and property professionals treat the on-line data source with confidence, procedures were adopted to ascertain obvious errors and steps were taken to rectify them.

According to the case summary of the *Original Raw Data* dataset, shown in Table 8.2 data were missing in four variables. These were *Price, Lot size, Yield* and *Tenure*. The eighty missing cases were identical in the *Price, Lot Size* and the *Yield* variables. The reason for this is that the missing cases in all of these variables were the result of missing *Price* values since both *Lot Size* and *Yield* are a function of *Price*. Twenty of the twenty-two missing *Tenure* values were to be found in the eighty missing cases appertaining to *Price*.

Table 8.2 Case summary of *Original Raw Data* dataset for *Region*, *Lot Size*, *Banking Company*, *Year* and *Yield (%)*

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Region	1012	100.0%	0	.0%	1012	100.0%
Lot Size	932	92.1%	80	7.9%	1012	100.0%
Banking Co	1012	100%	0	0%	1012	100.0%
Year	1012	100.0%	0	.0%	1012	100.0%
Yield (%)	932	92.1%	80	7.9%	1012	100.0%

Removing the eighty cases with missing *Price* values left 932 cases remaining in the *Price Included* dataset. Those eighty cases excluded by price realization included premises let to most of the banking companies, including the three with most cases, and premises located in all regions. They included the fifteen Yorkshire Bank premises offered at auction on 23rd May, 2000, but for which the hammer prices were kept out of the public domain. There was no given explanation for why the realization prices of the remaining lots were not recorded. It may have been that they were either sold prior to auction or that they failed to achieve their reserve prices.

Removing the eighty cases with missing *Price* data also excluded twenty of the cases with missing *Tenure* data. Thus, two cases with missing *Tenure* data remained in the *Price Included* dataset.

After re-examining the summary statistics for the *Price Included* dataset, there were found to be a number of banking companies represented in the dataset with less than five cases. Those banking companies were excluded at this stage by the creation of the *Low Counts Excluded* dataset. Once the cases excluded by either Price were removed, 909 cases remained in the *Low Counts Excluded* dataset. Amongst the twenty-three cases removed was one of the two remaining cases for which *Tenure* was unknown.

Since this research only covers freehold banking-halls, those cases comprising leasehold and heritable tenures were removed from the *Low Counts Excluded* dataset to create a new dataset known as the *Purified Freehold* dataset. Twenty-nine cases of leasehold and five cases of heritable premises were removed. The one remaining case with unknown *Tenure* was also removed, leaving a dataset comprising 874 cases.

8.2.3 Purified Data Set

The *Purified Freehold* dataset included all those cases comprising freehold, and in the instance of Scotland feudal and feuhold, premises for which rents reserved and hammer prices were recorded. Through the sequence of successive datasets from which cases were systematically removed, the *Purified Freehold* dataset omitted cases with either missing *Price* data, banks scoring counts of less than five, or leasehold premises. According to the case summary of the *Purified Freehold* dataset, as shown in Table 8.3, no data were missing in any of the cases for any of the main variables identified in qualitative study.

Table 8.3 Case summary of missing observations for each variable in the *Purified Freehold* dataset

Output of missing observations for each variable

Case No	Date	Bank	Region	Location	Price	Rent	Yield
0	0	0	0	0	0	0	0
Lot Size	**Tenure**	**Quarter**	**Bank No**	**Region Code**	**Size Code**	**Tenure Code**	**Year**
0	0	0	0	0	0	0	0

8.2.4 Discussion on the data review and purification

Before the start of Exploratory Data Analysis, the *Original Raw Dataset* was carefully scrutinized for any obvious errors. For example, it was noted that the original sources of the data had erroneously placed a number of the banking-halls into the incorrect regions. In particular, a number of premises situated in the South-west had been wrongly recorded as being in Wales. After exhaustive inspection, these errors were corrected in the dataset. The data comprised within the *Original Raw Dataset* was purified by the systematic removal of cases with missing data. This was achieved by the creation of a number of intermediate datasets from which those cases with the missing data were removed. Once this had been undertaken, those forms of tenure other than freehold were removed in order to arrive at the *Purified Freehold* dataset. The sequence of datasets through which this was achieved is shown in Table 8.1. Further checks on the quality of the *Purified Freehold Dataset* were undertaken during the Exploratory Data Analysis where box-plots were used to check for erroneous outliers in the data (see section 8.6).

.

8.3 Summary Statistics for the *Purified Freehold Dataset*

8.3.1 Yield Distribution

These summary statistics for the *Purified Freehold* dataset provide the following:

- A histogram of *Yield (%)* showing the bi-modal distribution of the dependent variable
- Bar charts for the categorical variables
- Observations in the conclusions about the bi-modal attribute of the dependent variable and about the lack of balance in the observed data

The histogram reproduced in Figure 8.1 illustrates the distribution of *Yield (%)* with respect to the *Purified Freehold* dataset. Against the histogram is plotted a curve, which not only shows that the distribution is very clearly bi-modal, but also enables the plotting of co-ordinates at the trough of the curve between the two modes. It is a probability density curve which gives an estimate of the underlying distribution of *Yield (%)*. The reading at the bottom of the trough between the two modes with respect to the *Purified Freehold* dataset was shown to be 6.46 per cent. The data fall into two groups with the lower group containing slightly more than half of the observations.

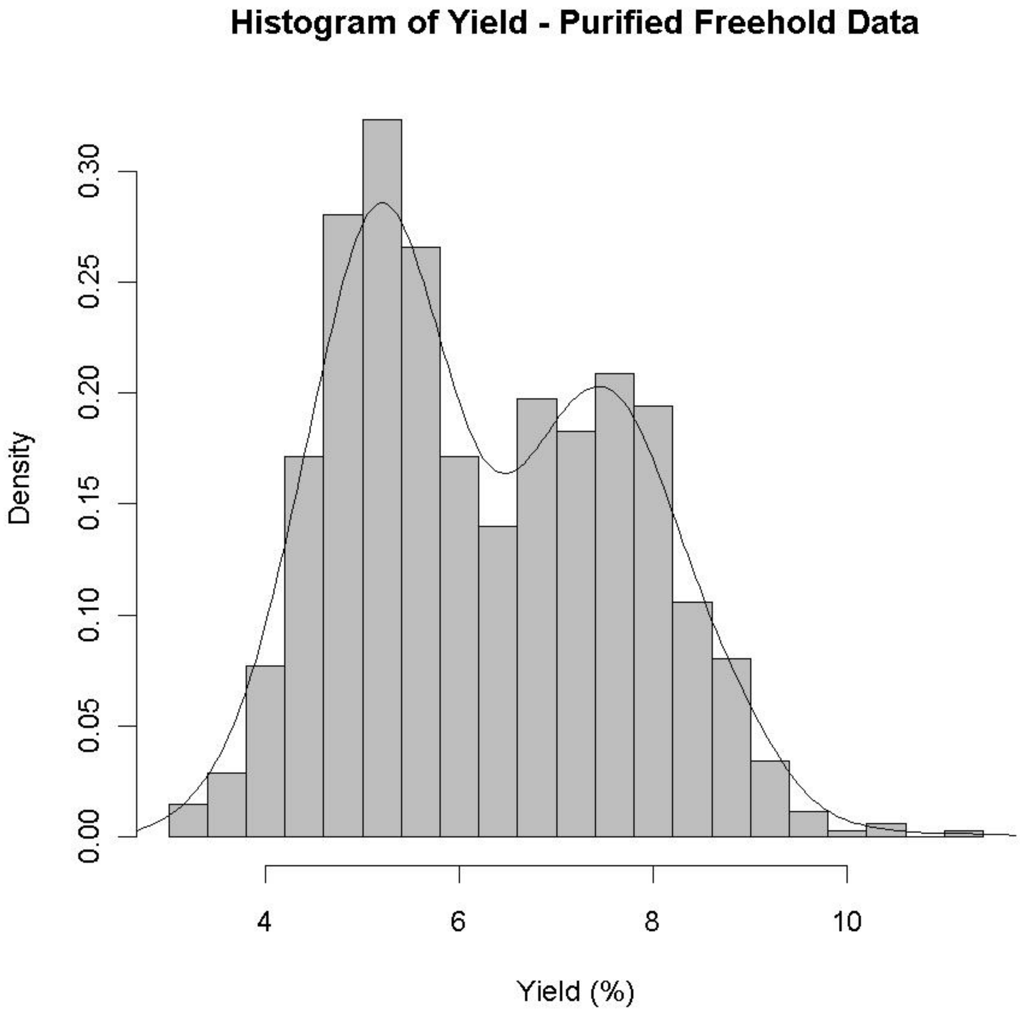


Figure 8.1 Histogram of densities per *Yield (%)* in the *Purified Freehold* dataset

8.3.2 Counts for each banking company

The *Purified Freehold* dataset retained twelve categories of banking company. The number of counts per banking company ranged from 5 to 339. As the bar-chart in Figure 8.2 shows, the number of counts by banking company was concentrated in a small number of categories. At the lower end, both Halifax Bank and Royal Bank of Scotland (RBS) each only scored five counts. At the upper end of the score of counts, three banking companies predominated. These three were Barclays, HSBC and Lloyds TSB. The combined counts of these three scored 691. The next highest three only scored counts of 48, 42 and 38 respectively. The cases comprising the three main banking companies accounted for 79.1 per cent of all cases.

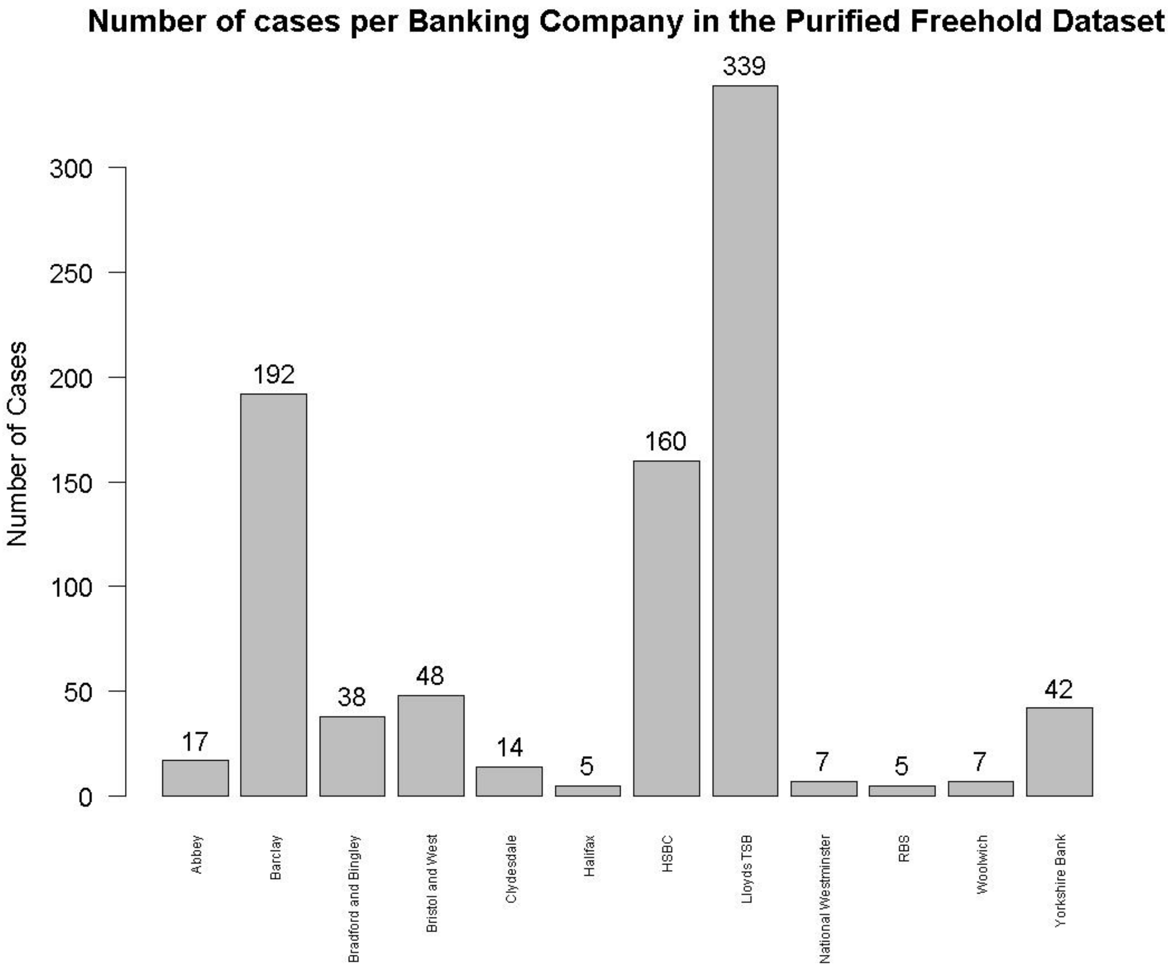


Figure 8.2 Bar chart of frequencies per *Banking Company* in the *Purified Freehold* dataset

8.3.3 Counts for each region

The *Purified Freehold* dataset retained ten categories of region. The number of counts per region ranged from 32 to 232. The counts per region are shown in Figure 8.3. The largest number of counts was in the South-East and the South-West of England. East Anglia and Scotland scored the lowest counts.

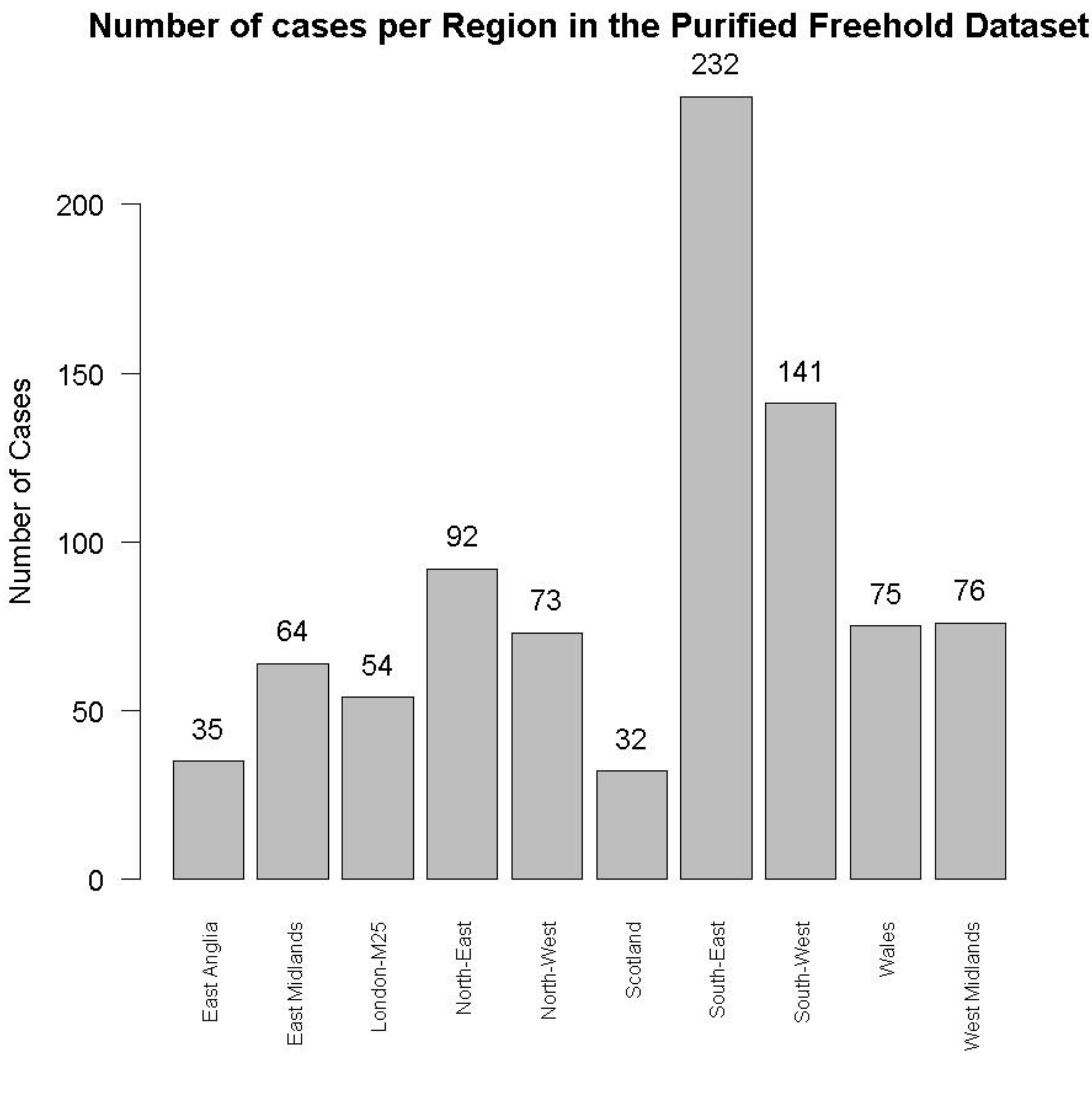


Figure 8.3 Bar chart of frequencies per *Region* in the *Purified Freehold* dataset

8.3.4 Counts for each lot size

The *Purified Freehold* dataset comprised three categories of lot size. The large lot size category contained 419 cases which equated with 47.9 per cent of the cases. The medium lot size comprised more cases than the small lot size category. Combined these two categories comprised 52.1 per cent of the cases. Hence, the small and medium categories combined contained a very similar number of cases to the large category singularly. The counts per lot size are shown in Figure 8.4.

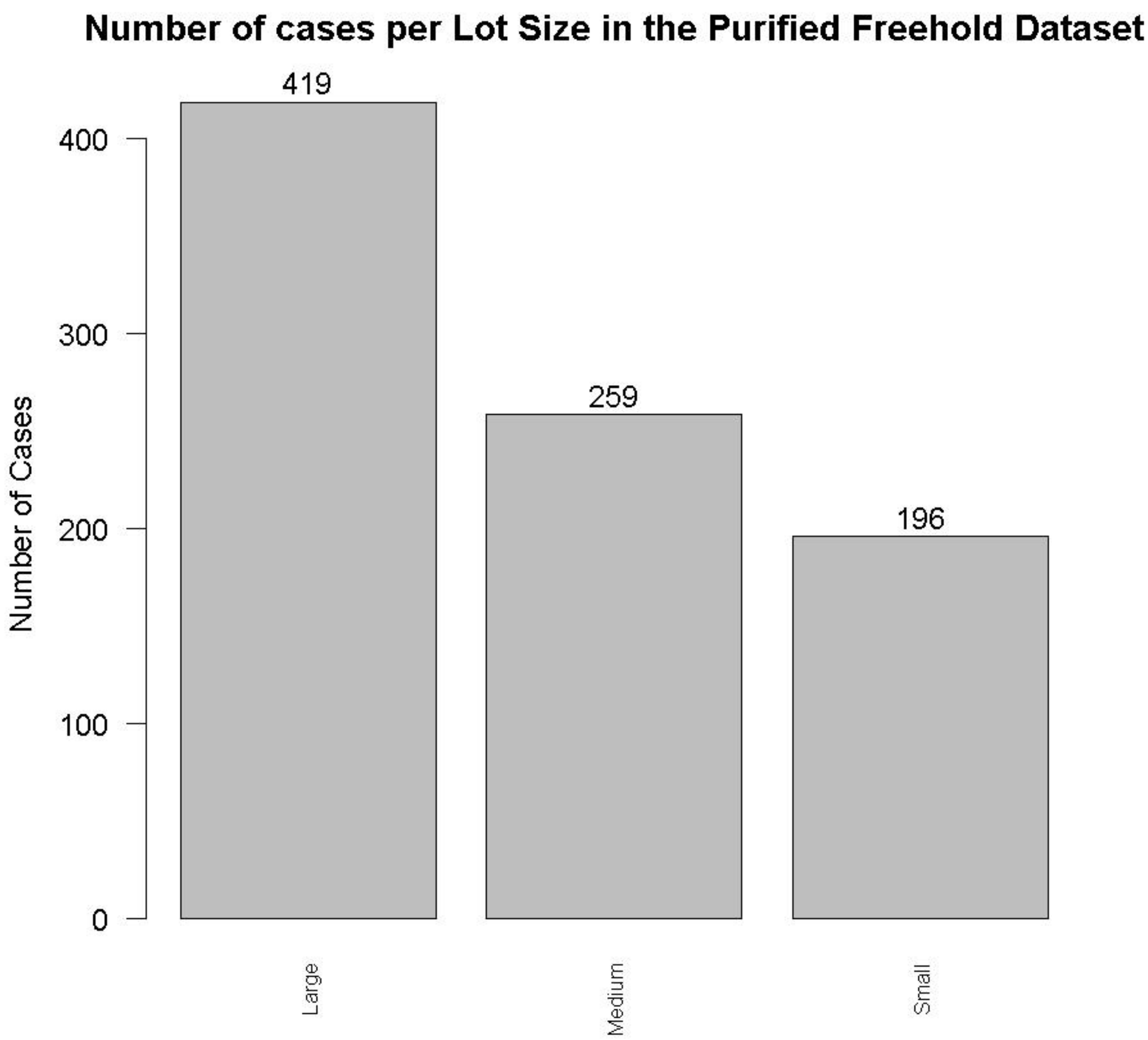


Figure 8.4 Bar chart of frequencies per *Lot Size* in the *Purified Freehold* dataset

8.3.5 Counts for year

The *Purified Freehold* dataset separately retained all eight years for which data were available to the study. The largest number of counts was in the years 2001 and 2006, which scored 196 and 221 respectively. The lowest count was in 2000, which only scored 26. The counts per year are shown in Figure 8.5.

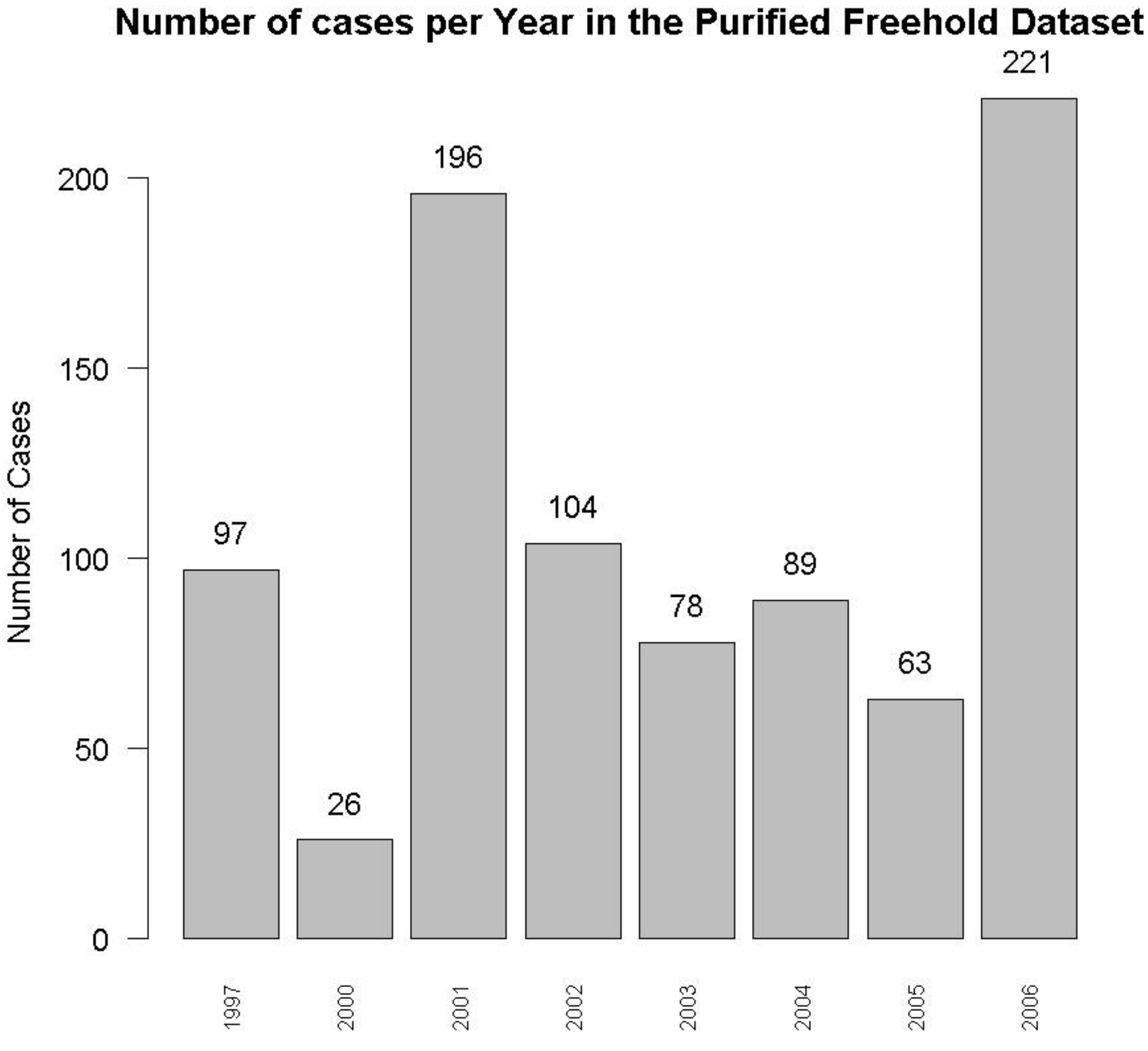


Figure 8.5 Bar chart of frequencies per *Year* in the *Purified Freehold* dataset

8.4 Cross-tabulations

8.4.1 Introduction to cross-tabulation

Cross tabulations are an effective and revealing way of summarising large datasets and of identifying the relationships between the variables. Cross-tabulations give a visual representation of the multivariate probability distribution defining the relationship between the variables.

Cross-tabulations require all the variables to be categorical. That is to say that they must be expressed as factors. Cross-tabulations present the data in a matrix called a contingency table. Each cell shows the number of cases for a given combination of factors.

All the variables in the *Purified Freehold* dataset, except *Yield,* can be treated as factors, which is the term usually used for categories in statistical analyses. Hence, *Yield* needed to be converted into a categorical form in order for it to be capable of being included in cross-tabulations. The histogram of *Yield* with respect to this dataset clearly showed it to be divided into two groups with a bi-modal distribution. Therefore, in order to be able to undertake analysis that would give meaningful results, it was appropriate to convert *Yield* into a categorical variable with two factor levels. Converting *Yield* into two factor levels produced a contingency table comprising 5,760 cells (i.e. 2 x 8 x 10 x12 x 3 = 5,760 possible combinations of factor levels). However, the *Purified Freehold* dataset only comprised 874 observations. Therefore, most of the cells in the contingency table remained empty. Thus, arranged in this way the data were sparse.

With so many cells in the contingency tables, it was extremely difficult to present the data in a form which would give a clear idea of the joint distribution of observations across five variables. Also, with so many unpopulated cells, it was not possible to undertake any meaningful form of contingency table analysis due to the data being so sparse (Agresti, 1996: 28, 34 and 194).

In order to visualize and interpret data and to perform contingency table analysis, the number of cells in the contingency table needed to be sufficiently reduced. Such reduction is often referred to as *collapse*. Furthermore, Agresti states that contingency tables need to average at least five observations per cell in order to facilitate valid analysis. Thus with 874 cases, the maximum number of cells that could be permitted with respect to the *Purified Freehold* dataset was 175.

8.4.2 Number of levels for each of the factors

In its uncollapsed form, the *Purified Freehold* dataset comprised five variables, including the dependent variable, each containing a differing number of categories. Multiplied together, these variables and their factor levels produced 5,760 cells in the contingency table. The levels for each variable and the total number of cells that they produced in the contingency table are shown in Table 8.4.

Table 8.4 Number of levels for each variable in the *Purified Freehold* dataset

Variable	Number of factor levels
Region	10
Lot Size	3
Banking Company	12
Year	8
Yield	2
Number of combinations	**5,760**

The sparseness and lack of balance in the *Purified Freehold* dataset necessitated compaction of the data. One way of compacting the data is to reduce the number of predictor variables. Agresti (1996: 190) states that sparseness is common where there are a large number of variables or factor levels. This can make the analysis less reliable. However, the qualitative study gave almost equal credence to the four predictors selected: *Region*, *Lot Size*, *Banking Company* and *Year*. The qualitative study, underpinned by the study in theoretical perspectives, has made the selection of these predictors robust. In order to maintain robustness, these predictors needed to be retained in some form. However, in order to enable analysis, these variables needed to be collapsed. Such collapse was achieved by combining two or more categories, or factor levels, within any given variable. The result was that the collapsed variables comprised fewer categories. This in turn meant that the number of cells within the contingency table was reduced. The intention of doing this was try to create a new dataset that was less sparse and more balanced.

8.4.3 Yield distribution within the *Three Main Banks* datasets

The production of a histogram of *Yield* as a continuous variable, very graphically illustrated the distribution of the *Three Main Banks* datasets. This is reproduced in Figure 8.6. Here the Y-axis is the probability density. The histogram very clearly shows that the spread of the data is bi-modal. By plotting a curve to this histogram, it is easy to see where the trough lies. The R statistical package did allow the co-ordinates of the trough to be read. This reading showed that the bottom of the trough was 6.35 per cent. That is to say that the lowest count of cases between the two modes lies at 6.35 per cent.

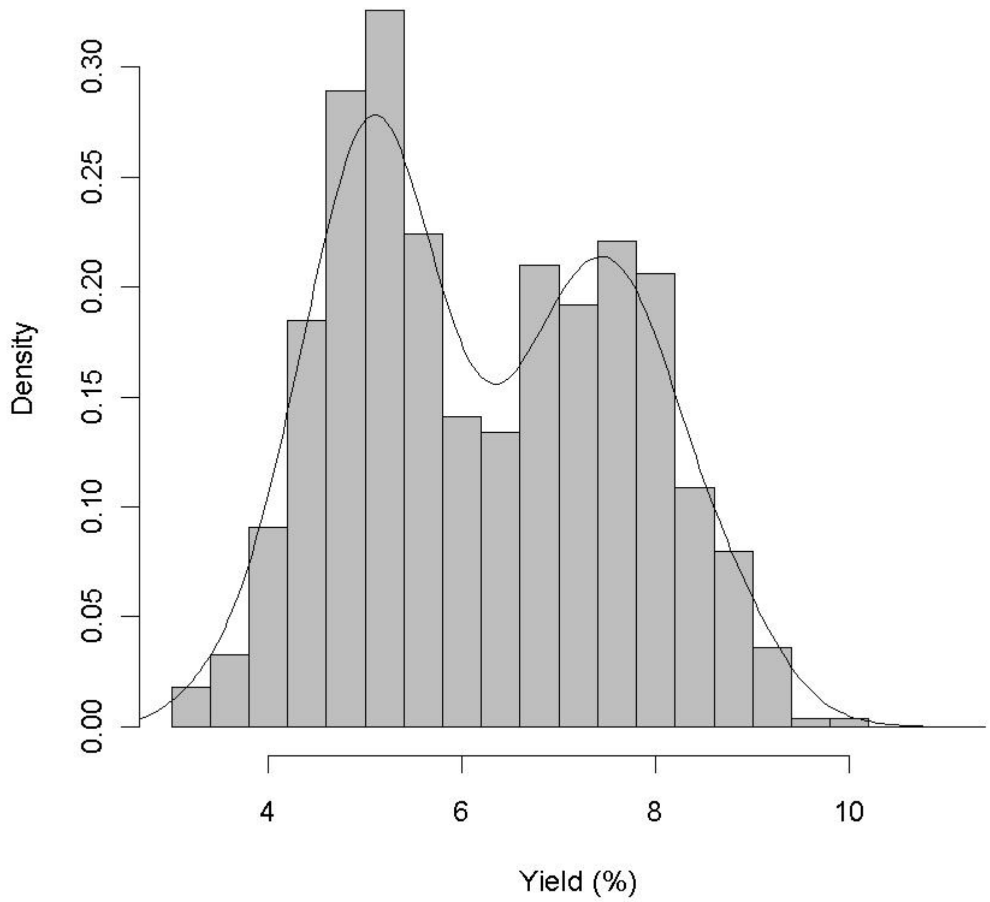


Figure 8.6 Histograms of densities per Yield (%) in the Three Main Banks dataset

Figure 8.6 clearly shows that the distribution of frequencies of *Yield (%)* is not normal. It is bi-modal. The histogram relates to the freehold premises of the three main banks over the entire period of the study. It produces two peaks. When a histogram with increments of 0.25 per cent on the X axis was produced, it was easy to see in which 0.25 per cent increments the two peaks lies. This more complex histogram is reproduced in Figure IV.1 in Appendix IV where it is shown as a frequency distribution. It could easily be seen from that figure that the value of the first of the two modes fell between 5.00% and 5.25%. The value of the second mode lies between 7.25% and 7.50%.

Table 8.5 gives the descriptive statistics of freehold premises sold by the three main banks by *Yield (%)*. It shows the lowest value of *Yield* as being 3.17% and the highest value as being 9.90%. Therefore, the low point of the probability density curve of the *Three Main Banks* datasets is 6.35172%. Hence, it seemed appropriate to adopt 6.35%, as the value on the scale coinciding almost precisely with both the median of the datasets and as the bottom of the trough between the two modes within them, as being the dividing point if *Yield* is to be further collapsed into just two categories. Such a division into two categories also gave an almost even split in the number of frequencies. A frequency table was produced on the basis of a split at the 6.35% increment. This showed a split of 372 (53.8%) and 319 (46.2%) between low yields and high yields respectively. This is shown in Table 8.6.

Table 8.5 Descriptive statistics of freehold premises sold by the three main banks by Yield (%)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Yield (%)	691	3.17	9.90	6.2229	1.44680
Valid N (listwise)	691				

Table 8.6 Frequency table of freehold premises sold by the three main banks by Yield Group

Yield Group

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Low	372	53.8	53.8	53.8
	High	319	46.2	46.2	100.0
	Total	691	100.0	100.0	

Since the exploratory data analysis showed that the dependent variable was bi-modal, final collapsing of the dependent variable in the *Three Main Banks* datasets for the building of the model was in the form of two categories rather than an alternative number of factor levels. During Exploratory Data Analysis, *Yield* was, after being transformed into a categorical variable, collapsed at different factor levels. Once the data had been fully examined, it became clear that the spread of the dependent variable was very distinctly bi-modal.

Therefore, *Yield Group* was selected as the final form of the dependent variable when it was collapsed from *Yield*. Accordingly, *Yield Group* comprised two factor levels which were split so as to take account of the trough between the two modes. As can be ascertained from the Frequency Table of freehold premises sold by the three main banks by *Yield Group* (Table 8.6), the two resultant categories were almost evenly split. They were split between 53.8 per cent and 46.2 per cent per respective category.

8.4.4 Collapsing the data

On the basis of previous analysis, three banking companies had 79.1 per cent of the observations. So collapsing the dataset to only those three main banks retained most of the data, but reduced the number of banks from twelve to three and thus reduced the number of cells in the contingency table by a factor of four.

The bi-modal yield distribution suggested that yields fall naturally into two groups for the three main banks with a low yield group below 6.35 per cent and a high yield group at or above 6.35 per cent.

The calendar variable *Year* was treated as a factor rather than a continuous variable, having eight levels. These eight levels were 1997 and 2001 to 2006 inclusive. The variable was collapsed by replacing these eight levels with a new variable *Time*, which comprised two levels. The two levels of *Time* were defined as:

Early = 1997, 2000, 2001and 2002;

Late = 2003, 2004, 2005 and 2006.

Year was collapsed in the contingency table into *Time* in this way, because such collapse was representative of the yield compression that occurred during the study period.

By reducing the number of levels for the calendar factor from eight to two, the number of cells in the contingency table was reduced by a further factor of four.

The original data comprised a geographical factor, *Region*, with ten categories or factors. The geographical classification could sensibly be reduced, in several ways, from ten *Regions* to four *Provinces* by combining regions in different combinations. These *Provinces* equated with what the literature describes as *Super-regions*. This reduced the number of cells by a further factor of 2.5.

Finally, the *Lot Size* factor could be reduced from three levels to two by combining the 173 small and 197 medium levels into a new small 370 level. This collapsed variable was named *Lot Size Group*. This further reduced the cell count in the contingency table by a factor of 1.5. When *Lot Size* comprised three factor levels, the low factor level did not have any counts in the last year of the study period. Therefore, when these three factor levels were collapsed into two factor levels, combining the original low and medium factor levels seemed to be the convenient way of doing it.

Overall, with 3, 2, 2, 4 and 2 levels respectively for the factors, *Bank, Yield, Time, Provinces* and *Lot Size Group*, this produced a contingency table with 691 observations in ninety-six cells. This was based upon 3 x 2 x 2 x 4 x 2 = 96 cells. The levels for each variable and the total number of cells that they produced in the contingency table are shown in Table 8.7. A five-dimensional cross tabulation with ninety-six cells was much easier to interpret and gave a much clearer idea of the distribution of observations across the five variables. With an average of more than seven observations per cell, the data in this form was also potentially suitable for multivariate contingency table analysis using log-linear modelling.

Table 8.7 Number of levels for each variable in the *Three Main Banks B* and the *Three Main Banks C* datasets

Variable	Number of factor levels
Province	4
Lot Size Group	2
Banking Company	3
Time (Early or Late Years)	2
Yield Group	2
Number of combinations	**96**

8.4.5 Collapsing regions into provinces

There was no unambiguously best way of collapsing the ten regions into four super-regions or provinces. These were collapsed on the basis of what seemed a sensible and meaningful way based upon geographic proximity or regional similarity. Two different ways were adopted.

In order to reduce the *Purified Freehold* data to the two respective datasets in which *Region* is collapsed to the two different versions of *Province*, an intermediate dataset was created. This intermediate dataset was named *Three Main Banks A*. In this dataset all those cases relating to premises occupied by banking brands other than the three main ones by number were removed. Hence, the *Three Main Banks A* dataset only retained cases comprising freehold premises occupied by Barclays, HSBC or Lloyds TSB. This dataset comprised 691 cases. It facilitated the creation of both the *Three Main Banks B* and the *Three Main Banks C* datasets through collapsing *Region* into the respective version of *Province* in these datasets. Accordingly, the dataset *Three Main Banks B* has *Region* collapsed as *Provinces A* and the dataset *Three Main Banks C* has *Region* collapsed as *Provinces B*. The respective forms of collapse of *Region* are also shown in a simplified format in Tables 8.8 and 8.9 respectively.

Table 8.8 ***Region* collapsed as *Provinces A***

Regions	**Collapsed to Provinces A**
London-M25 South-East South-West	South
East Anglia West Midlands East Midlands	Midlands
North-East North-West	North
Wales Scotland	Celtic

Table 8.9 ***Region* collapsed as *Provinces B***

Regions	**Collapsed to Provinces B**
London-M25 South-East	London & South-East
South-West Wales	Wales & South-West
East Anglia West Midlands East Midlands	Midlands
North-East North-West Scotland	North Britain

The *Provinces A* and the *Provinces B* datasets produced the same data summaries for each of the variables with the exception of the *Provinces* variable. The *Provinces* variable produced different data summaries, because the variable *Region* was collapsed into *Province* in different ways for the two datasets. The data summaries for all the input variables are reproduced as follows in Tables 8.10, 8.11, 8.12, 8.13, 8.14, 8.15 and 8.16.

Table 8.10 **Summary of input data of both datasets by *Banking Company***

Banking Company	**Number of cases**
Barclays	192
HSBC	160
Lloyds TSB	339
Total	**691**

Table 8.11 **Summary of input data of both datasets by *Region***

Region	**Number of cases**
East Anglia	32
East Midlands	53
London-M25	49
North-East	61
North-West	53
Scotland	15
South-East	193
South-West	103
Wales	65
West Midlands	67
Total	**691**

Table 8.12 **Summary of input data of both datasets by *Lot Size Group***

Lot Size Group	**Number of cases**
Small	370
Large	321
Total	**691**

Table 8.13 **Summary of input data of both datasets by *Time***

Time	**Number of cases**
Early	391
Late	300
Total	**691**

Table 8.14 **Summary of input data of both datasets by *Yield***

Yield	**Number of cases**
Low	372
High	319
Total	**691**

Table 8.15 Summary of input data of the *Provinces A* dataset by *Province*

Province	Number of cases
North of England	114
Midlands	152
South of England	345
Celtic Regions	80
Total	**691**

Table 8.16 Summary of input data of the *Provinces B* dataset by *Province*

Province	Number of cases
London & South-East	242
Wales & South-West	168
Midlands	152
North Britain	129
Total	**691**

The frequency tables show that the counts by province were more evenly spread with respect to *Provinces B* than they were for *Provinces A*.

8.4.6 Cross-tabulation tables

With the data collapsed as described in the previous sections, the cross-tabulation contingency tables now comprised ninety-six cells. With the data in this form, meaningful cross tabulation could be performed. The cross-tabulation for *Provinces A* is shown in Table 8.17. The cross-tabulation for *Provinces B* is shown in Table V.1 in Appendix V.

Table 8.17 Crosstabulation of the *Three Main Banks B* dataset (Provinces A)

	Yield	Bank	Time	Lot Size	Provinces A	Frequency
1	Low	Barclays	Early	Small	North	0
2	High	Barclays	Early	Small	North	0
3	Low	HSBC	Early	Small	North	0
4	High	HSBC	Early	Small	North	24
5	Low	Lloyds TSB	Early	Small	North	4
6	High	Lloyds TSB	Early	Small	North	19
7	Low	Barclays	Late	Small	North	0
8	High	Barclays	Late	Small	North	0
9	Low	HSBC	Late	Small	North	2
10	High	HSBC	Late	Small	North	0
11	Low	Lloyds TSB	Late	Small	North	3
12	High	Lloyds TSB	Late	Small	North	2
13	Low	Barclays	Early	Large	North	0
14	High	Barclays	Early	Large	North	2
15	Low	HSBC	Early	Large	North	1
16	High	HSBC	Early	Large	North	0
17	Low	Lloyds TSB	Early	Large	North	0
18	High	Lloyds TSB	Early	Large	North	13
19	Low	Barclays	Late	Large	North	30
20	High	Barclays	Late	Large	North	0
21	Low	HSBC	Late	Large	North	8
22	High	HSBC	Late	Large	North	0
23	Low	Lloyds TSB	Late	Large	North	4
24	High	Lloyds TSB	Late	Large	North	2
25	Low	Barclays	Early	Small	Midlands	0
26	High	Barclays	Early	Small	Midlands	0
27	Low	HSBC	Early	Small	Midlands	0
28	High	HSBC	Early	Small	Midlands	9
29	Low	Lloyds TSB	Early	Small	Midlands	17
30	High	Lloyds TSB	Early	Small	Midlands	37
31	Low	Barclays	Late	Small	Midlands	3
32	High	Barclays	Late	Small	Midlands	0
33	Low	HSBC	Late	Small	Midlands	1
34	High	HSBC	Late	Small	Midlands	0
35	Low	Lloyds TSB	Late	Small	Midlands	8
36	High	Lloyds TSB	Late	Small	Midlands	0
37	Low	Barclays	Early	Large	Midlands	0
38	High	Barclays	Early	Large	Midlands	0
39	Low	HSBC	Early	Large	Midlands	0
40	High	HSBC	Early	Large	Midlands	0
41	Low	Lloyds TSB	Early	Large	Midlands	3
42	High	Lloyds TSB	Early	Large	Midlands	15
43	Low	Barclays	Late	Large	Midlands	39
44	High	Barclays	Late	Large	Midlands	0
45	Low	HSBC	Late	Large	Midlands	10
46	High	HSBC	Late	Large	Midlands	0

47	Low	Lloyds TSB	Late	Large	Midlands	7
48	High	Lloyds TSB	Late	Large	Midlands	3
49	Low	Barclays	Early	Small	South	0
50	High	Barclays	Early	Small	South	0
51	Low	HSBC	Early	Small	South	6
52	High	HSBC	Early	Small	South	47
53	Low	Lloyds TSB	Early	Small	South	40
54	High	Lloyds TSB	Early	Small	South	47
55	Low	Barclays	Late	Small	South	12
56	High	Barclays	Late	Small	South	0
57	Low	HSBC	Late	Small	South	6
58	High	HSBC	Late	Small	South	0
59	Low	Lloyds TSB	Late	Small	South	19
60	High	Lloyds TSB	Late	Small	South	0
61	Low	Barclays	Early	Large	South	0
62	High	Barclays	Early	Large	South	2
63	Low	HSBC	Early	Large	South	0
64	High	HSBC	Early	Large	South	2
65	Low	Lloyds TSB	Early	Large	South	12
66	High	Lloyds TSB	Early	Large	South	24
67	Low	Barclays	Late	Large	South	98
68	High	Barclays	Late	Large	South	1
69	Low	HSBC	Late	Large	South	11
70	High	HSBC	Late	Large	South	2
71	Low	Lloyds TSB	Late	Large	South	15
72	High	Lloyds TSB	Late	Large	South	1
73	Low	Barclays	Early	Small	Celtic	0
74	High	Barclays	Early	Small	Celtic	0
75	Low	HSBC	Early	Small	Celtic	0
76	High	HSBC	Early	Small	Celtic	26
77	Low	Lloyds TSB	Early	Small	Celtic	2
78	High	Lloyds TSB	Early	Small	Celtic	31
79	Low	Barclays	Late	Small	Celtic	1
80	High	Barclays	Late	Small	Celtic	0
81	Low	HSBC	Late	Small	Celtic	1
82	High	HSBC	Late	Small	Celtic	0
83	Low	Lloyds TSB	Late	Small	Celtic	2
84	High	Lloyds TSB	Late	Small	Celtic	1
85	Low	Barclays	Early	Large	Celtic	0
86	High	Barclays	Early	Large	Celtic	0
87	Low	HSBC	Early	Large	Celtic	0
88	High	HSBC	Early	Large	Celtic	1
89	Low	Lloyds TSB	Early	Large	Celtic	0
90	High	Lloyds TSB	Early	Large	Celtic	7
91	Low	Barclays	Late	Large	Celtic	4
92	High	Barclays	Late	Large	Celtic	0
93	Low	HSBC	Late	Large	Celtic	3
94	High	HSBC	Late	Large	Celtic	0
95	Low	Lloyds TSB	Late	Large	Celtic	0
96	High	Lloyds TSB	Late	Large	Celtic	1

The cross-tabulations give an indication of where the highest counts lie and where there remain empty cells.

With respect to *Provinces A*, forty-two, nearly forty-four per cent, of the cells remained empty. This cross-tabulation produced eleven cells with counts of twenty-four or more. Due to imbalance that remained in the data even after they had been collapsed to ninety-six cells in the contingency table, there were noticeable clusters of high counts. Of the three main banks retained, it was noticeable that within the highest counts, HSBC appeared to be concentrated within the early *Time* period, whereas Barclays appeared to be concentrated within the late *Time* period. Those HSBC cases in the highest counts tended to be both in the smaller *Lot Group* and in the early *Time* period. On the other hand, those Barclays lots that were within the highest counts were in both the larger *Lot Group* and the late *Time* period, but had a broad geographical spread. Specific banking companies undertook more sale-and-leaseback transactions in given years than others.

The results from a chi-square test of independence between the factors in the *Provinces A* data are given in Table 8.18. The result is significant at beyond the 0.1% level, indicating that the factors are not independent. That is to say that the value of at least one variable is strongly dependent on the value of one or more of the other variables. Thus, there are clearly significant relationships between the variables which were worthy of further investigation.

Table 8.18 Chi-square test *Provinces A* cross-tabulation

```
Number of cases in table: 691
Number of factors: 5
Test for independence of all factors:
        Chisq = 2230.2, df = 87, p-value = 0
        Chi-squared approximation may be incorrect
```

With respect to *Provinces B*, forty of the cells remained empty, which was two less than for *Provinces A*. This cross-tabulation also produced eleven cells with counts of twenty-four or more. Some geographical bias seemed apparent from the cross-tabulation with HSBC scoring very low counts in Wales and the South-West, but better in London and the South-

East. HSBC also had one relatively higher score in the North of Britain factor level within *Provinces B*.

The results from a chi-square test of independence between the factors in the *Provinces B* data are given in Table 8.19. The result is significant at beyond the 0.1% level, indicating that the factors are not independent. Thus, again there are clearly significant relationships between the variables which were worthy of further investigation.

Table 8.19 Chi-square test *Provinces B* cross-tabulation

```
Number of cases in table: 691
Number of factors: 5
Test for independence of all factors:
        Chisq = 2304.5, df = 87, p-value = 0
        Chi-squared approximation may be incorrect
```

Two-way cross-tabulation tables between *Yield Group* and each of the other variables are shown in Tables 8.20, 8.21, 8.22, 8.23 and 8.24. With respect to the collapsed variables *Provinces A* and *Provinces B*, both are shown in respective two-way cross-tabulation tables with *Yield Group*. These marginal two-way tables satisfy the conditions of having no zero entries and having less than twenty per cent of the cells with counts of less than five. Thus the corresponding contingency tables are suitable for log-linear analysis. The cell with the lowest count, which scores five, was in Table 8.23: *Yield*Bank*. All the other cells in all the other tables score more than five.

Table 8.20 Two-way Table of Yield Group and Provinces A

Yield Group * Provinces A Crosstabulation

Count

		Provinces A				
		North	Midlands	South	Celt	Total
Yield Group	1	52	88	219	13	372
	2	62	64	126	67	319
	Total	114	152	345	80	691

Table 8.21 **Two-way Table of Yield Group and Provinces B**

Yield Group * Provinces B Crosstabulation

Count

		Provinces B				
		London & SE	Wales & SW	Midlands	North Britain	Total
Yield Group	1	163	69	88	52	372
	2	79	99	64	77	319
	Total	242	168	152	129	691

Table 8.22 **Two-way Table of Yield Group and Lot Size Group**

Yield Group * Lot Group Crosstabulation

Count

		Lot Group		
		Small	Large	Total
Yield Group	1	127	245	372
	2	243	76	319
	Total	370	321	691

Table 8.23 **Two-way Table of Yield Group and Banking Company**

Yield Group * Banking Co Crosstabulation

Count

		Banking Co			
		HSBC	Lloyds TSB	Barclays	Total
Yield Group	1	49	136	187	372
	2	111	203	5	319
	Total	160	339	192	691

Table 8.24 **Two-way Table of Yield Group and Time**

Yield Group * Time Crosstabulation

Count

		Time		
		Early	Late	Total
Yield Group	1	85	287	372
	2	306	13	319
	Total	391	300	691

Two-way cross-tabulation tables between *Region* and *Province* are shown respectively with respect to the *Provinces A* and *Provinces B* datasets in Tables V.2 and V.3 both in Appendix V. These are shown to give some further illustration of how the data populate the cells.

8.4.7 Discussion on cross-tabulations

Cross-tabulations are an easy way of looking at data in order to ascertain how the cells are filled. To enable cross-tabulation to do this, all the data must be categorical. This can mean transforming some numeric data into categorical data. Accordingly, such transformation was undertaken to ensure that all the data was categorical. The cross-tabulation showed that the data in the *Purified Freehold* dataset were very sparse and unbalanced. The cross-tabulation showed that there were many empty cells, and that few cells had high counts.

The hypothesis test suggests that the relationship of the factors is not independent. That is to say that at least one factor is influenced by at least one of the others. Thus, there is a relationship worthy of further investigation.

The two-way marginal tables average more than five entries per cell, have no zero entries and no more than twenty per cent of the cells have counts of less than five (Agresti, 1996: 194). The contingency tables are therefore suitable for modelling using log-linear analysis

(Agresti, 1996: 190-194). Such analysis allows investigation of the relationships suggested in the previous paragraph.

8.5 Line plots

8.5.1 Number of premises sold by each of the three main banks by year

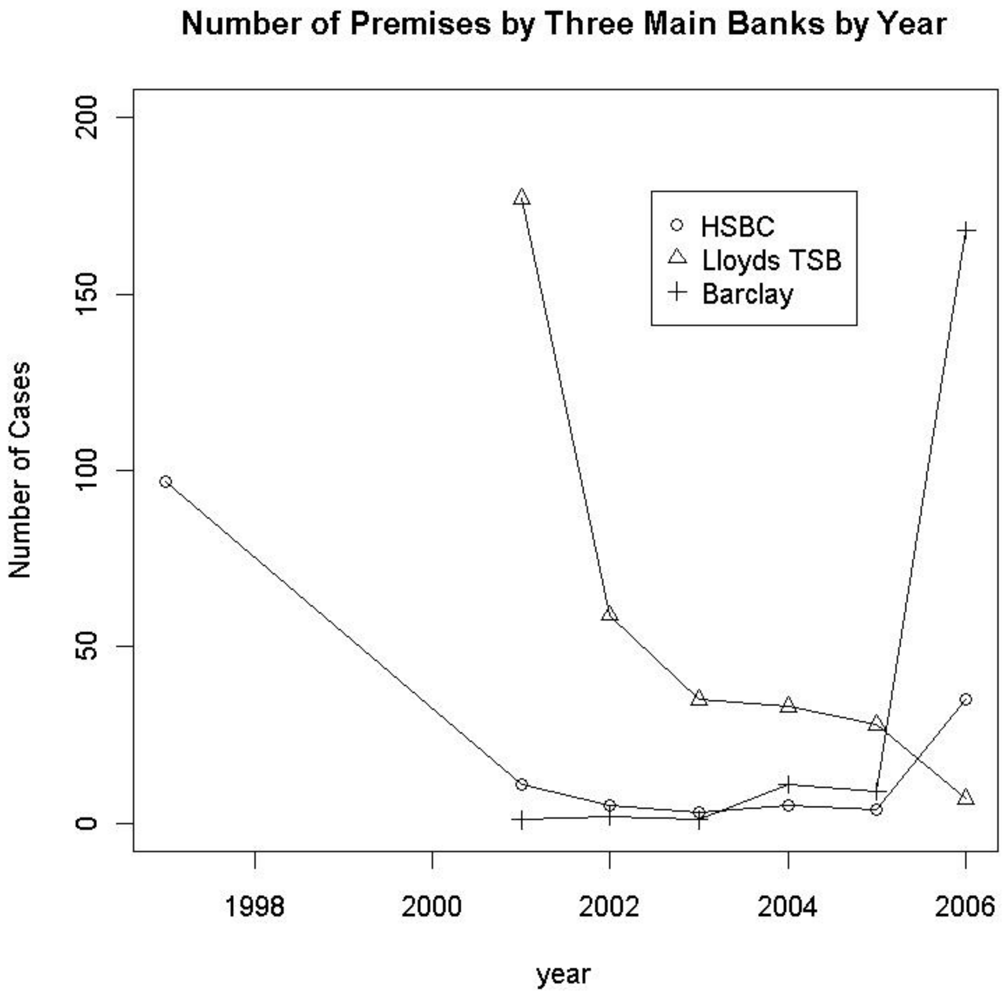


Figure 8.7 Number of Premises by Three Main Banks by Year

The lineplot in Figure 8.7 represents the number of premises sold by each of the three main banks by year. This showed that HSBC sold most of its lots during the beginning of the study period, whereas Barclays sold most of its lots during the end of the period and Lloyds TSB spread its disposal more evenly across the period, but with a peak in the middle of the period. This gave a clear graphical representation of the imbalance of disposals by banks over time.

8.5.2 Yield for the Three Main Banks by Year

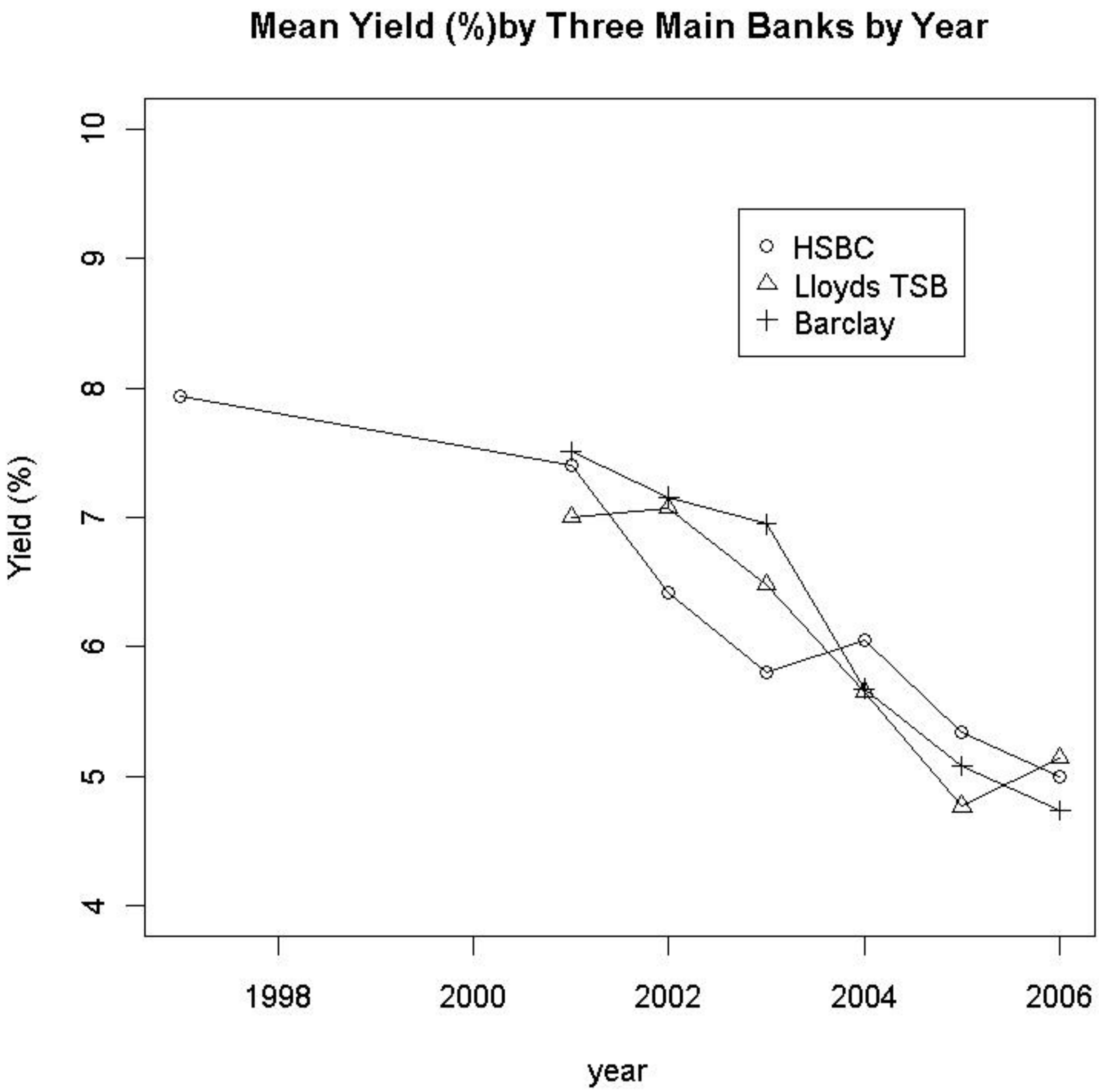


Figure 8.8 Mean Yield (%) by Three Main Banks by Year

There is no clear separation between the three lines shown in Figure 8.8 (i.e. the lines cross one another). Hence, there is no obvious clear-cut difference between banking companies. The lineplot was very suggestive of a yield effect over time. Further analysis was required to ascertain if these suggestions could be substantiated.

8.5.3 Yield for the Three Many Banks by Lot Size over Years

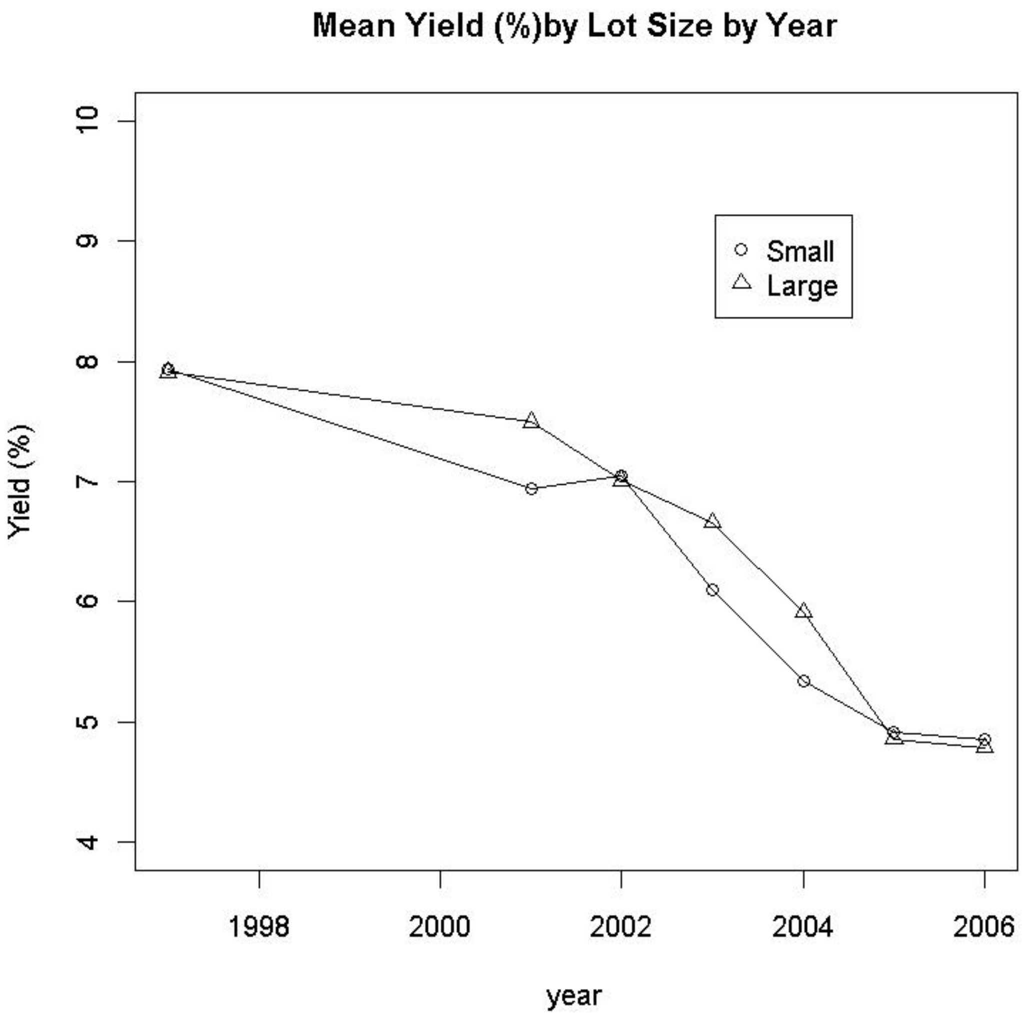


Figure 8.9 Mean Yield (%) by Lot size by Year

The lineplot in Figure 8.9 shows a tenuous suggestion of a *Lot Size* effect, which is as expected based on the findings of the qualitative study (see section 6.4.7). The lineplot shows that the mean yields by lot size by year converged in the years 1997 and 2002.

Otherwise, the small lot sizes generally generated lower mean yields than the larger ones for the bulk of the study period with the exception of the years 2005 and 2006.

8.5.4 Yield for the Three Many Banks by Provinces over Years

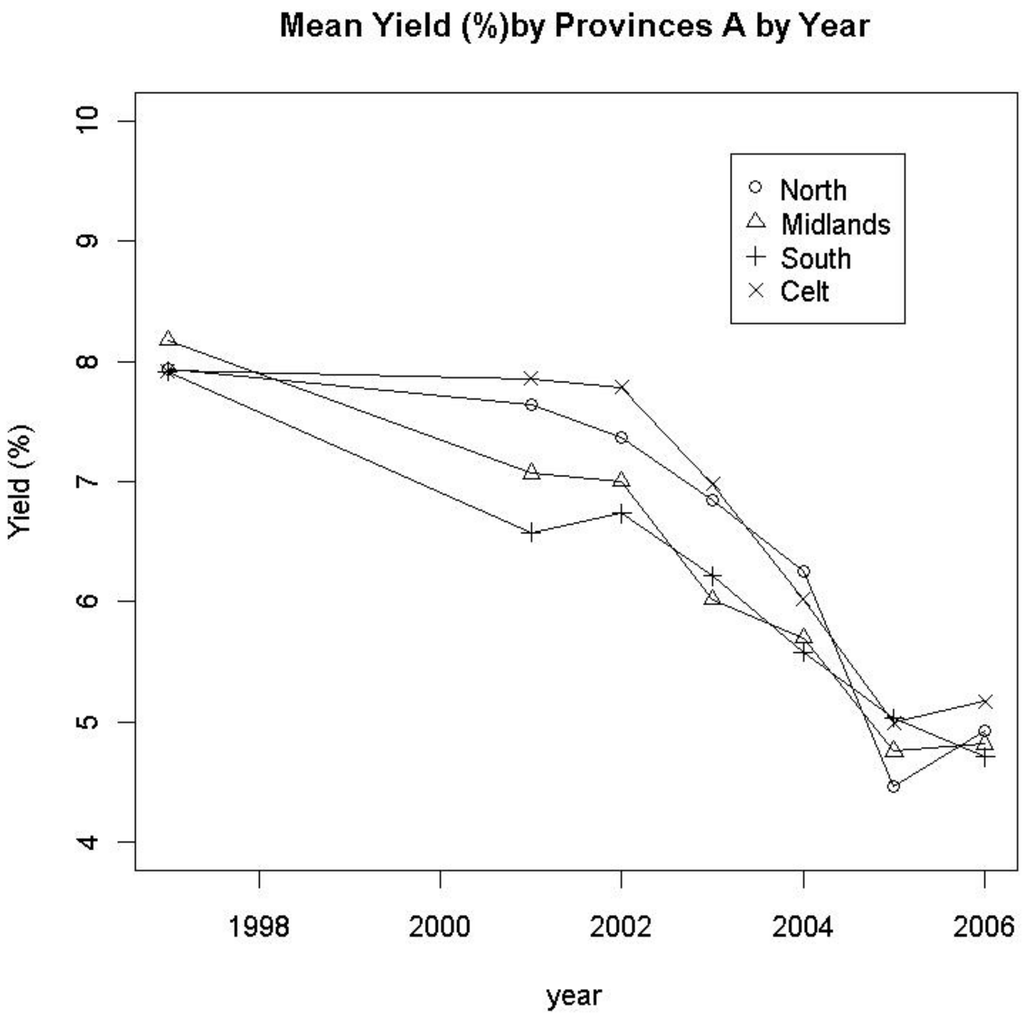


Figure 8.10 Mean Yield (%) by Provinces A by Year

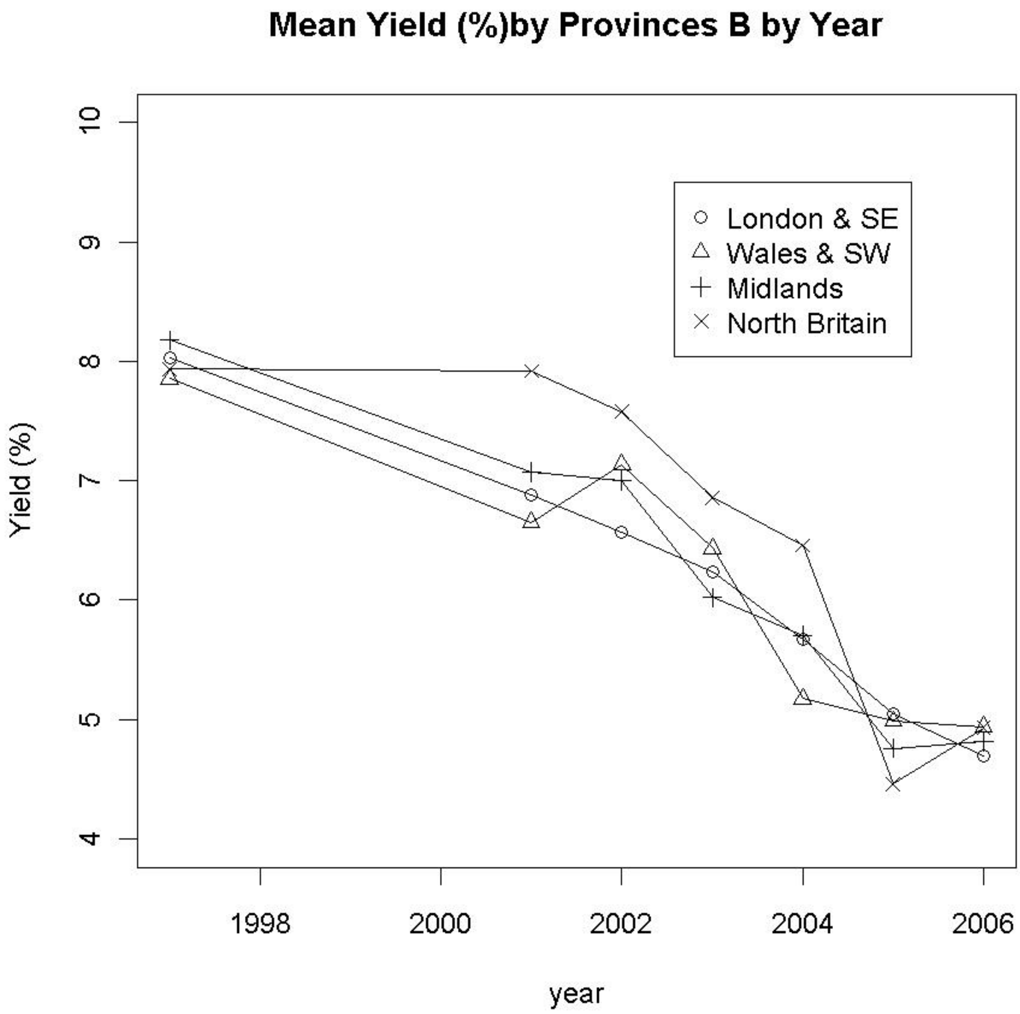


Figure 8.11 **Mean Yield (%) by Provinces B by Year**

The two lineplots with respect to both *Provinces A* and *Provinces B* in Figures 8.10 and 8.11 show a tenuous suggestion of a *Province* effect. This was expected, since *Province* is a collapsed form of *Region* and *Region* was identified in the qualitative study (Chapter 6) as likely to have an effect.

8.5.5 Discussion on lineplots

The lineplots helped to confirm the imbalance in the data even when it had been collapsed to the *Three Main Banks* datasets. It was shown by lineplot that even premises let to the three main banks were not sold at the same rate over time during the period subject to the study. In particular, a concentration of HSBC premises was sold at the early part of the period, whereas a concentration of Barclays premises were sold at the end of the period. No particular *Bank* effect was identified by lineplot, since the lines crossed one another. A tenuous suggestion of *Lot Size* effect was shown by lineplot. A tenuous *Province* effect was shown by lineplot.

8.6 Boxplots

8.6.1 Introduction to boxplots

Boxplots, sometimes known as Box-whisker diagrams, give the scores within the data simple graphical representation of the median, spread and inter-quartile range of a variable (Field, 2005). The box-whisker diagrams incorporate the following attributes:

- The inter-quartile range, which is the middle fifty per cent of the distribution, is denoted by the box.
- The heavy line within the box denotes the median and its position within the box shows whether the distribution is symmetrical or skewed (Field).
- Notches on the sides of the boxes represent the confidence interval in the median and a lack of overlap between the notches indicates a significant difference between the medians of the respective categories Chambers *et al.* (1983: 62).
- The whiskers, which are the bars at the end of the vertical lines, represent the largest and the smallest values in a sample which cannot be considered outliers. Any outliers may need to be checked to ascertain that they are not corrupted or incorrectly input data.

According to Chambers *et al.* (1983: 62), there is substantial evidence that the medians of two boxes differ if their respective notches do not overlap. During this study, notches were generated on the sides of the boxplots with a view to ascertaining if the medians of the respective categories differed.

Boxplots can be a useful way of identifying errors in the recording of the data insofar as they highlight outliers. Such plotting was undertaken with respect to the *Three Main Banks A,* and *C* datasets. In each boxplot, the outliers were noted and checked. Once having been confirmed as being correct, these outlier scores were left in the dataset for the analysis.

8.6.2 Distribution of Yield by Region

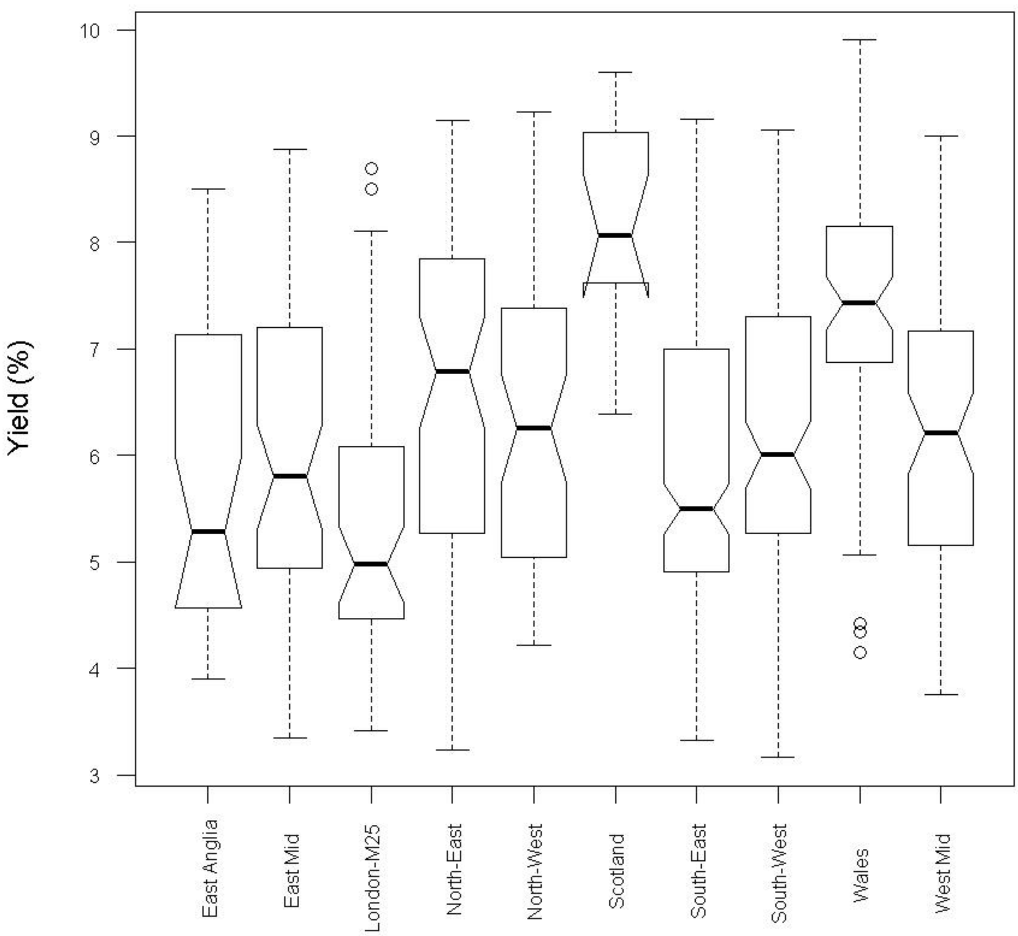


Figure 8.12 Boxplot of Yield (%) by Region

The few outliers in the boxplots of Yield by Region for the three main banking companies were identified as the tail ends of the distributions. The lack of the overlap of the notches in the boxplot suggested that *Region* did have a significant effect on *Yield*. However, since the cross-tabulations had shown that the data lacked balance, caution had to be exercised before that conclusion could be accepted.

8.6.3 Distribution of Yield by Lot Size

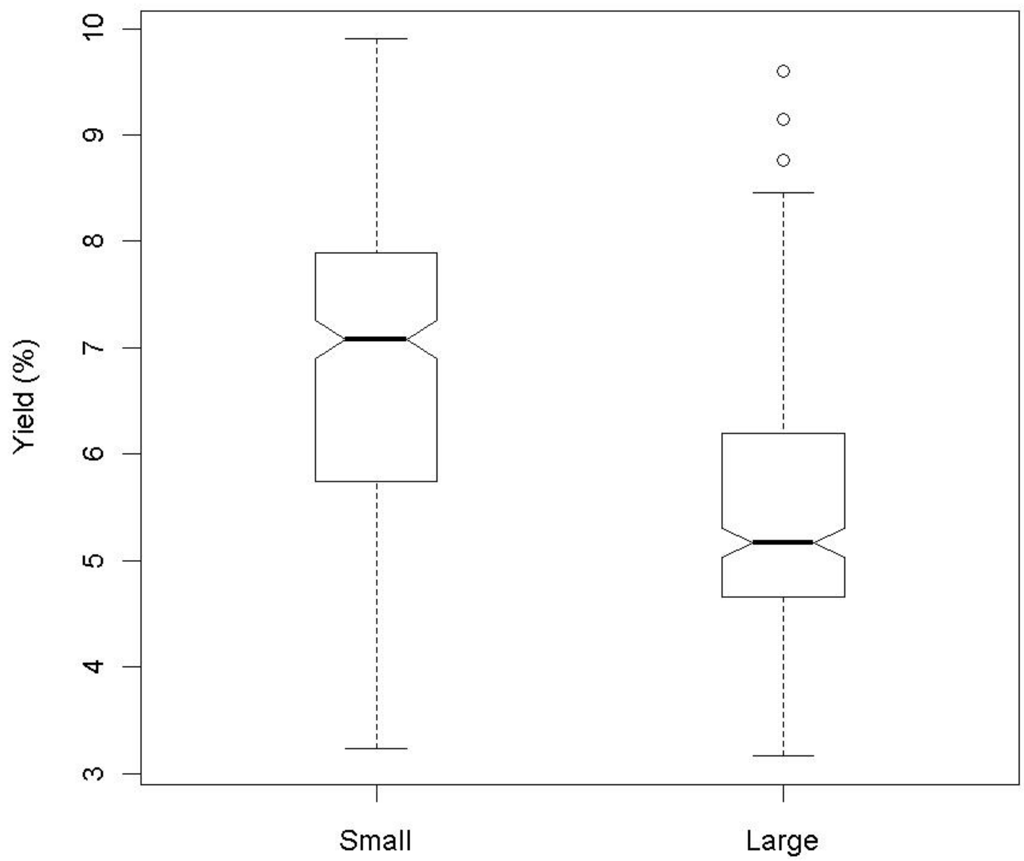


Figure 8.13 Boxplot of Yield (%) by Lot Size

The three outliers in the large lot size in the boxplot of Yield by Lot Size were identified as being the tail end of the distribution. The lack of overlap in the notches in the boxplot suggested a significant difference arising from *Lot Size*, but this had to be qualified. Before this suggestion was accepted, there was a need to establish that this was not due to the effect of time with smaller lots being sold at the beginning of the period and larger lots being sold at the end of the period.

8.6.4 Distribution of Yield by Banking Company

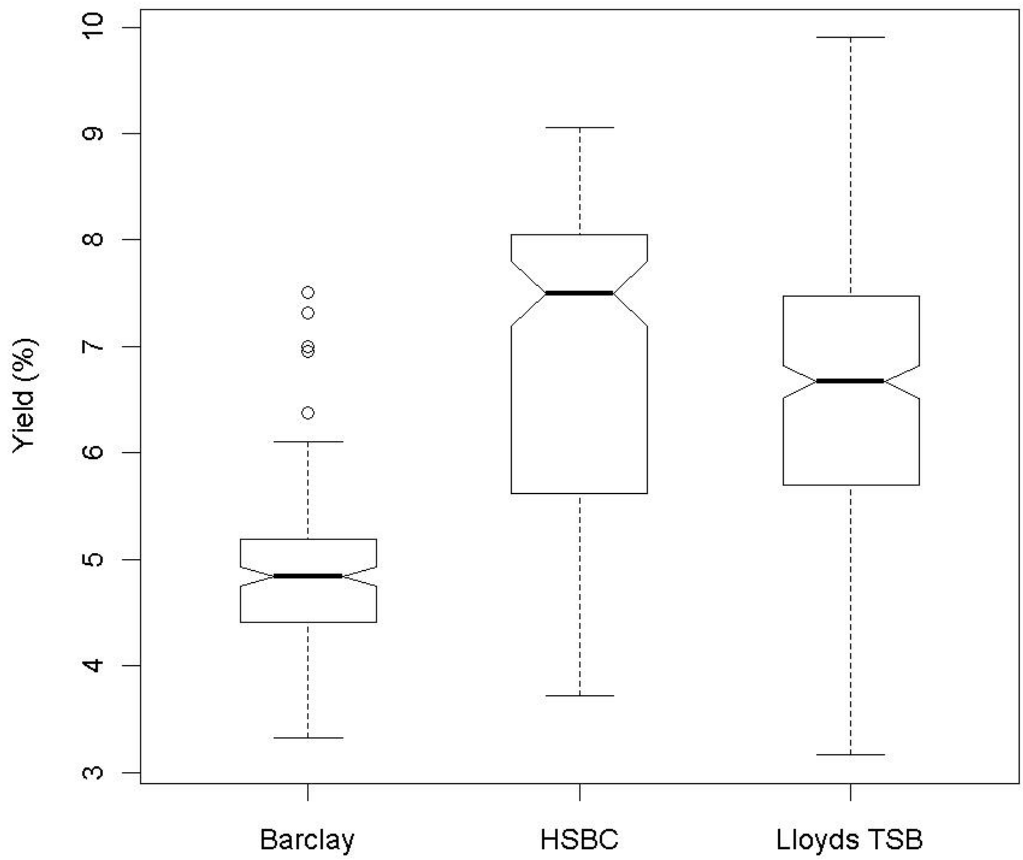


Figure 8.14 Boxplot of Yield (%) by Banking Company

The few outliers in the boxplot were identified as the tail ends of the distributions. The notches for the three categories did not all overlap. This implies that a significant difference exists between banking companies. It is necessary to be cautious about the interpretation, because most of the premises sold at the beginning of the period were sale-and-leaseback properties sold by HSBC and most of the premises sold at the end of the period were sold by Barclays, whereas the sale-and-leaseback properties sold by Lloyds were more evenly spread over the period, but with a peak in the middle of the period. However, it could be that other influences may be buried in the data.

8.6.5 Distribution of Yield by Year

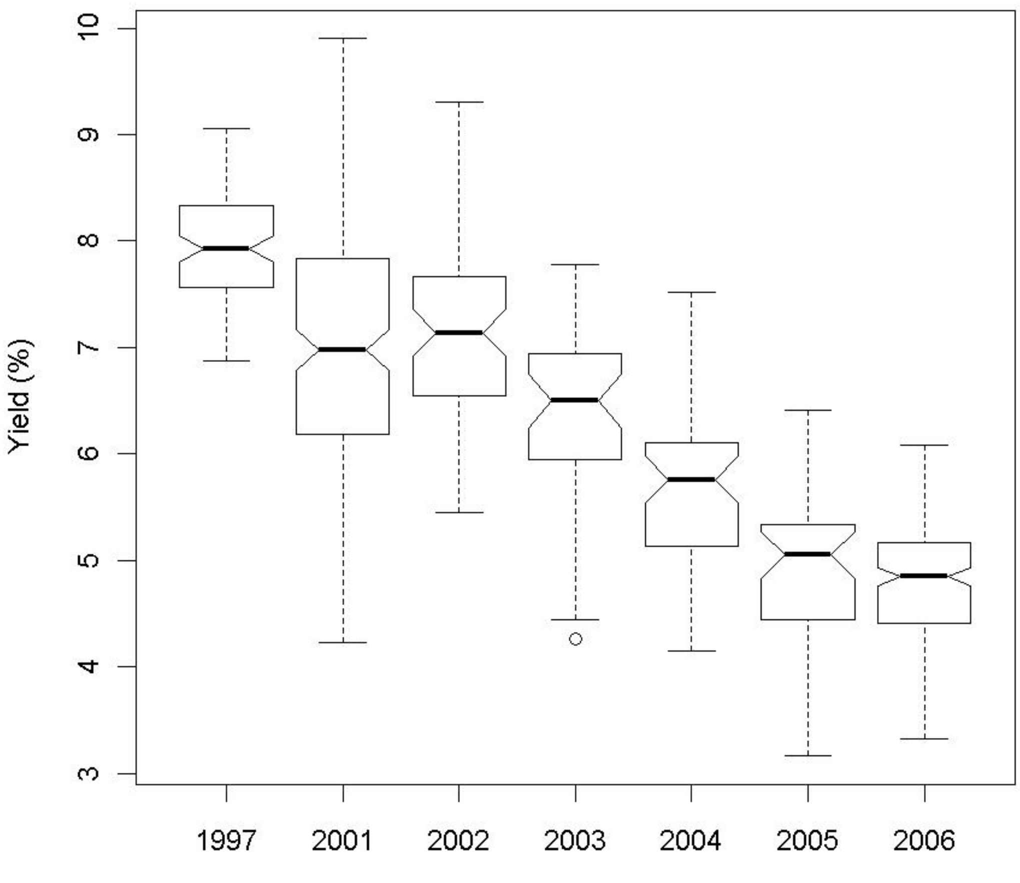


Figure 8.15 Boxplot of Yield (%) by Year

There was only one outlier in the boxplot of Yield by Year. The outlier was checked to establish that the data were not erroneous. It was identified as being the tail end of the distribution and was therefore retained in the analysis. The boxplot for yield by year between 1997 and 2006 inclusive showed that yields fell significantly during that period.
This fall was in accordance with a general contraction of the yields of investment properties in the United Kingdom during the study period (Investment Property Databank, 2009). However, this situation was complicated by other influences buried in the data. For example, some banking companies sold a much larger number of properties in certain years than in other years. Moreover, this differed from bank to bank. As the lineplot of the count of the three main banks sold over the years 2001 to 2006 illustrates in Figure 8.7, Lloyds TSB sold larger a number of properties in 2001 and 2002, whereas Barclays sold most of its properties in 2006.

8.6.6 Distribution of Yield by Time

The *Time* variable, which is a collapsed form *Year*, was found to contain more outliers in the boxplot. Since the *Year* variable had been found to contain only one outlier and since the data in that outlier were checked and found to be correct, the additional outliers found in the *Time* boxplot were accepted as comprising correct data. The occurrence of additional outliers in *Time* is an artefact of the way in which years were combined into early and late *Time*.

The separation between the notches for the two factor levels very strongly suggests that the factor *Time* is highly significant. This clearly denotes yield compression over the study period.

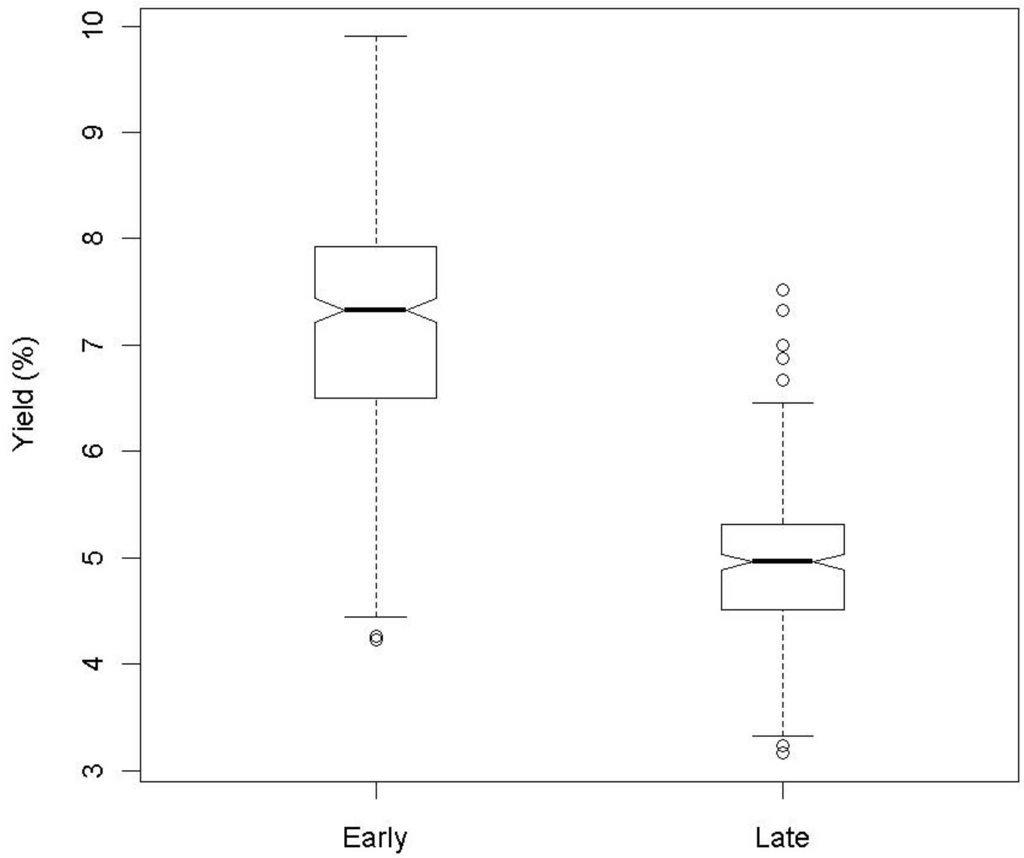


Figure 8.16 **Boxplot of Yield (%) by Time (Early and Late Years)**

8.6.7 Distribution of Yield by *Province*

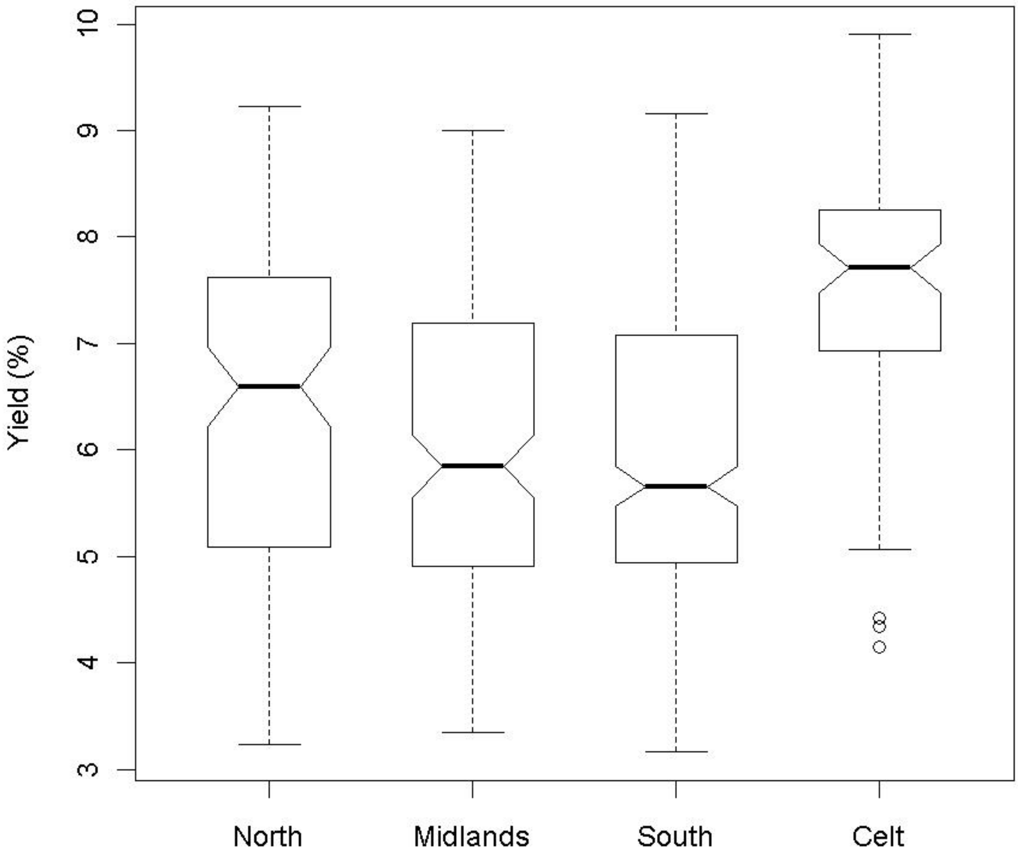


Figure 8.17 Boxplot of Yield (%) by Provinces A

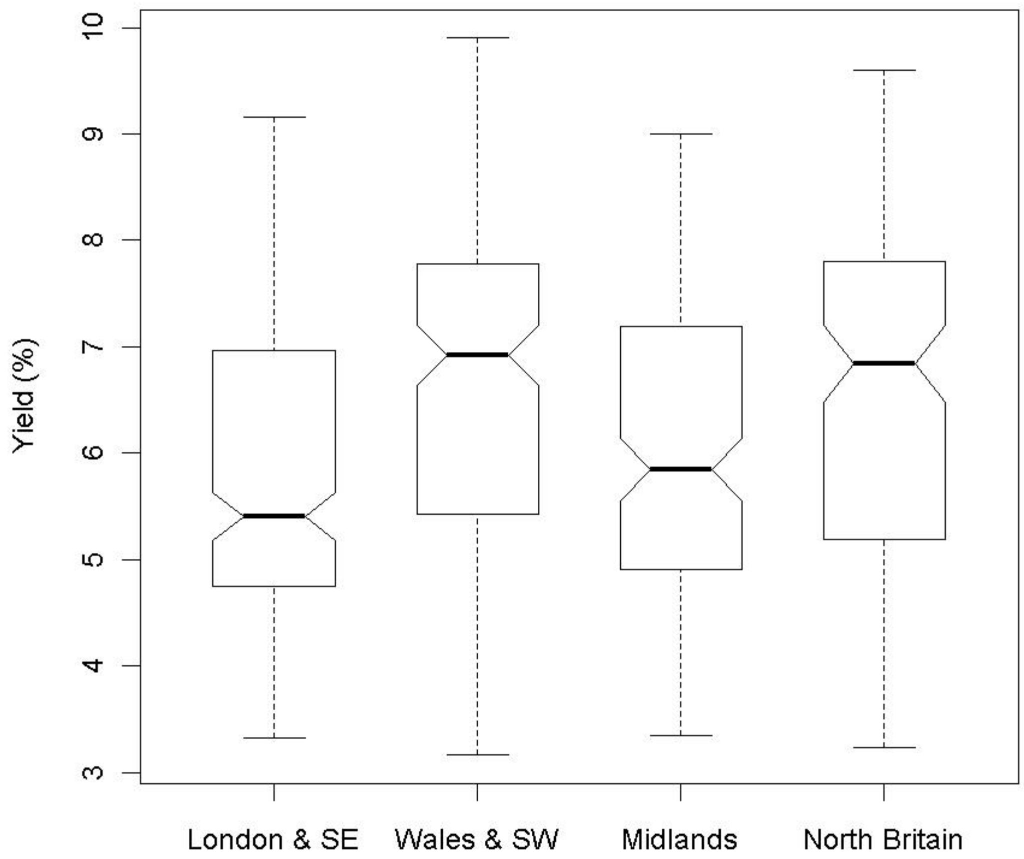


Figure 8.18 **Boxplot of Yield (%) by Provinces B**

Regions had been collapsed as two different forms of *Province*. These were represented by *Provinces A* and *Provinces B* respectively. No outliers were produced for *Provinces B*. *Provinces A* produced three outliers. These three outliers were in the Celt province and were all identified as being from the Wales region.

The notches in the plots for both provinces do not overlap. However, the absence of overlap is more pronounced in *Provinces B* than in *Provinces A*. Thus the boxplots show a significant difference in the factor levels within both forms of province. However, the difference between the factor levels is greater in *Provinces B*.

The boxplots suggest that *Provinces B* is better in having both few outliers and more significant in having greater differences in the notches in the plots.

8.6.8 Discussion on the boxplots

The main conclusions to be drawn from the boxplots were:

1. A small number of outliers were identified, all of which were the tail ends of the distributions. The data were all found to be valid and therefore there was no reason to remove any of these outliers.
2. All factors seemed to be significant, but care had to be taken with interpretation, since the data were unbalanced.
3. The boxplots suggested that *Yield* declined over time between 2001 and 2006. This effect was complicated by other factors buried in the data. However, the trend in yield contraction over the timespan in the research was known from other sources outside this study.

8.7 Tests of normality and homogeneity of variance

Tests of normality and homogeneity of variance were undertaken with respect to both the *Three Main Banks B* and the *Three Main Banks C* datasets. In both instances, all of those cases shown in the respective cross-tabulation tables to have counts of twenty-four or more were subjected to the tests. By coincidence, the respective cross-tabulation tables identified both datasets as having eleven cases with counts of twenty-four or more.

Both datasets were subjected to the Shapiro-Wilk normality test. According to Field (2006: 93), this test can be used to show if the distribution of a group of data departs from that of groups of like data having a normal distribution. A significant *p*-value does show a deviation from normality (Field, 2006: 744). According to Royston (1995), a Shapiro-Wilk statistic *p*-value < 0.1 should be considered an adequate level of normality.

Both datasets were then subjected to two tests of homogeneity of variance. These were the Fligner-Killeen test and the Bartlett test.

The Bartlett test of sphericity is an old test of homogeneity of variance, but it is sensitive to departures from normality (Bartlett, 1937). It checks that group variances are alike, but its practical value has been questioned (Field).

The Fligner-Killeen test relies upon the chi-square (x^2) and is a test of the null hypothesis that the variances of each of the groups are the same. According to Conover *et al.* (1981), this median has been shown in a simulation study to be one of the most robust tests of homogeneity of variance.

8.7.1 Tests of normality and homogeneity of variance with respect to *Provinces A*

The cross-tabulation for the *Three Main Banks B* dataset, comprising *Provinces A*, had eleven cells between them containing 443 of the total number of 691 cells. These amounted to sixty-four per cent of all the data.

The Shapiro-Wilk normality test produced the output reproduced in FigureVI.1 in Appendix VI. Three out of the eleven groups subjected to the Shapiro-Wilk test were shown to depart from normality at or beyond the 10% level. The largest group in this dataset was Group 9. The results for Group 9 and the results for the Bartlett test and the Fligner-Killeen test are reproduced in Figure 8.19

The Bartlett sphericity test produced a *p*-value = 0.000594 which is significant beyond the 0.1% level.

The Fligner-Killeen median test produced a *p*-value = 0.01290 which is significant beyond the 5% level.

```
        Shapiro-Wilk normality test

data:  yields
W = 0.9839, p-value = 0.2762
```

```
        Fligner-Killeen test of homogeneity of variances

data:  testframe$yieldvals and testframe$group
Fligner-Killeen:med chi-squared = 22.4663, df = 10,
p-value = 0.01290
```

```
        Bartlett test of homogeneity of variances

data:  testframe$yieldvals and testframe$group
Bartlett's K-squared = 30.9677, df = 10,
p-value = 0.000594
```

Figure 8.19 Tests of normality and homogeneity of variance with respect to *Provinces A*

8.7.2 Tests of normality and homogeneity of variance with respect to *Provinces B*

The cross-tabulation for the *Three Main Banks C* dataset, comprising *Provinces B*, had eleven cells between them containing 410 of the total number of 691 cells. These amounted to fifty-nine per cent of all the data.

The Shapiro-Wilk normality test produced the output reproduced in Figure VI.2 in Appendix VI. Four out of the eleven groups subjected to the Shapiro-Wilk test were shown to depart from normality at or beyond the 10% level. The largest group in this dataset was Group 9. The results for Group 9 and the results for the Bartlett test and the Fligner-Killeen test are reproduced in Figure 8.20.

The Bartlett sphericity test produced a *p*-value = 0.000025 which is significant at well beyond the 0.1% level.

The Fligner-Killeen median test produced a p-value = 0.000047 which is significant at well beyond the 0.1% level.

```
        Shapiro-Wilk normality test

data:  yields
W = 0.9856, p-value = 0.4685
```

```
        Fligner-Killeen test of homogeneity of variances

data:  testframe$yieldvals and testframe$group
Fligner-Killeen:med chi-squared = 37.4677, df = 10,
p-value = 4.696e-05
```

```
        Bartlett test of homogeneity of variances

data:  testframe$yieldvals and testframe$group
Bartlett's K-squared = 39.0853, df = 10,
p-value = 2.453e-05
```

Figure 8.20 Tests of normality and homogeneity of variance with respect to *Provinces B*

8.7.3 Conclusions on the tests of normality and homogeneity of variance

The large degree of lack of constance of variance makes the interpretation of the ANOVAs for these data extremely difficult. There is even less constance of variance between the groups within the *Provinces A* dataset than between the groups in the *Provinces B* dataset.

8.8 Conclusions from the initial Exploratory Data Analysis

8.8.1 Review of the data

The *Original Raw* dataset comprised 1,012 cases. The *Purified Freehold* dataset was arrived at after the systematic removal of certain cases. These cases were:

- Those with missing data
- Those with banking companies with scores less than five
- Those comprising leasehold or heritable tenures

Once those cases had been removed, 874 cases remained in the *Purified Freehold* dataset.

8.8.2 Examination of the *Purified Freehold* dataset

Investigation of the *Purified Freehold* dataset showed that *Yield* had a bi-modal distribution and fell naturally into two groups. Examination of the distribution of cases across the *Bank, Lot Size, Region* and Year variables showed that the data lacked balance. That is to say that there were a very unequal number of cases at factor levels.

8.8.3 Cross-tabulations

An initial attempt to produce a cross-tabulation of the *Purified Freehold* dataset showed that it was very sparse. There were only 874 cases to fill a cross-tabulation contingency table comprising 5,760 cells. To overcome this problem, the data were collapsed. The number of levels of each of the factors was reduced. Hence, the number of factor levels was broken down as follows:

- Only the three main banking companies with the highest counts were retained
- Geographically, ten regions were collapsed into four provinces
- Lot size was reduced from three factor levels in two in *Lot Size Group*
- The year factor was collapsed into two *Time* groups respectively comprising *early* (pre 2003) and *late* (2003 onwards)

This resulted in the *Three Main Banks B* and *C* datasets, each comprising 691 cases. In removing the lesser banks, only two-ninths of the data were lost. Thus, the truncated datasets comprising the data with respect to the freehold lots of the three main banks still retained seven-ninths of the data.

Examination of the cross-tabulation contingency table for the *Three Main Banks* datasets showed that although the data were still lacking balance and were still sparse, they could sensibly be tested for independence of the factors using a chi-square (χ^2) test.
This test produced a result significant at the 0.1 per cent level, clearly indicating that there was a relationship between the variables which required further investigation.

8.8.4 Lineplots

Lineplots of the data confirmed the lack of balance, even when it was collapsed to the three main banks, and gave a tenuous suggestion of a relationship between *Yield* and *Lot Size, Province* and *Time*. There was no obvious, simple relationship between *Yield* and *Bank*.

8.8.5 Boxplots

All the boxplots gave strong indications that *Yield* is influenced by the other four factors: *Bank, Lot Size, Time* and *Province*. These results had to be interpreted with great caution, however, because of the known lack of balance in the data.

8.8.6 Normality and homogeneity of variance

Tests of normality indicated that there were significant departures from normality in the datasets. Tests of homogeneity of variance showed very clearly that variance was not constant across the cells in the cross-tabulation contingency tables.

These two findings were important, because the assumption of normality and constant variance underpin Analysis of Variance (ANOVA). Thus, the results from any use of ANOVA to explore these data would need to be treated with great caution.

Additional analysis is required to check the utility of variance stabilizing transformations. Such checking could make an ANOVA model more robust (see section 9.2.1).

8.9 Log-Linear Analysis of the contingency table with yield as a categorical variable

8.9.1 Introduction

Log-linear analysis provides a method of ascertaining relationships between variables when the data are presented in the form of a contingency table (Agresti, 1996: p145). The existence of an interaction between two or more variables is a clear indication of a relationship between them inasmuch as the value of one variable is influenced by the value of another.

In log-linear analysis, the dependent variable, as shown in the model described in the following section, is the cell count in the frequency table. That is to say it is the number of observations per cell. The values of the categorical variables, also known as the factor levels, are used to explain the variation in counts from cell to cell (Agresti, 1996: 145).

Although yield does not appear as the dependent variable in the log-linear model, it is possible to conduct a log-linear analysis so that the emphasis is on revealing relationships between one particular variable and all the others. To do this, the analysis focuses specifically and only on the interactions which include the variable of particular interest. In this case, the variable of interest is *Yield.* All interactions between the other variables are regarded simply as describing the structure of the dataset being used to investigate *Yield.* As such, these interactions must not be removed from the model.

Interaction terms which include the variable of interest are regarded as representing relationships between this variable and the other explanatory variables. Weak or insignificant relationships are of little or no interest in terms of explaining the behaviour of the variable of interest. The model-fitting procedure therefore seeks to identify and reject these weak relationships whilst retaining only the stronger relationships with greater explanatory power.

Log-linear analysis relies only upon very weak assumptions about the form of the data. In particular, log-linear analysis does not require the data to be either normally distributed or to have constant variance. If the analysis, and in particular any inference drawn from it, is to be valid, then no observation should fall into more than one cell and the cells in the contingency table must be sufficiently well populated. Studies (Tabachnick and Fidell, 2001; Agresti, 1996: 28, 34 and 194) suggest that the sum of counts should be at least five times the number of cells. Furthermore, for all two-way associations, no more than twenty per cent of cells should have expected counts of less than five and no expected count should be less than one.

With *Yield* as a categorical variable, expressed as *Yield* Group, the cross-tabulation data from section 8.4 is a contingency table in the required form. Log-linear analysis was undertaken using both the *Three Main Banks B* and the *Three Main Banks C* datasets so that both the *Provinces A* and the *Provinces B* forms of collapse of *Region* could be used respectively.

8.9.2 The Model

Consider a two-way contingency table with I rows and J columns. Let the count in the (i,j)th cell be n_{ij}.

n_{11}	n_{12}	…	…	n_{1J}	$n_{1.}$
n_{21}	n_{21}	…	…	…	$n_{2.}$
…	…	…	…	…	…
…	…	…	…	…	…
n_{I1}	…	…	…	n_{IJ}	$n_{I.}$
$n_{.1}$	$n_{.2}$	…	…	$n_{.J}$	$n_{..}$

The marginal column totals are: $n_{.j} = \sum_i n_{ij}$

The marginal row totals are: $n_{i.} = \sum_j n_{ij}$

And the grand total is $n_{..}$ where $n_{..} = \sum_i n_{i.} = \sum_j n_{.j} = \sum_i \sum_j n_{ij}$.

Let μ_{ij} be the predicted value of the expected count in the (i,j)th cell.

The Null Model (the simplest model) is given by: $\mu_{ij} = m \quad \textit{for all } i, j$

where the maximum likelihood estimate of m is given by $m = \dfrac{n_{..}}{IJ}$

so that $\mu_{ij} = \dfrac{n_{..}}{IJ}$.

In **the Independent Effects Model** (i.e. assuming row and column effects are independent) the expected cell count is given by: $\mu_{ij} = n_{..}\pi_{ij}$

where π_{ij} is the proportion of counts in the (i,j)th cell,

and $\pi_{ij} = \pi_{i.}\,\pi_{.j}$

with $\pi_{i.}$ = the proportion of counts in row i

and $\pi_{.j}$ = the proportion of counts in column j.

Thus $\mu_{ij} = n_{..}\pi_{ij} = n_{..}\pi_{i.}\,\pi_{.j}$

The maximum likelihood estimates are $\pi_{i.} = \dfrac{n_{i.}}{n_{..}}$ and $\pi_{.j} = \dfrac{n_{.j}}{n_{..}}$.

Hence $\mu_{ij} = n_{..}\dfrac{n_{i.}}{n_{..}}\dfrac{n_{.j}}{n_{..}} = \left(\dfrac{n_{..}}{IJ}\right)\left(\dfrac{In_{i.}}{n_{..}}\right)\left(\dfrac{Jn_{.j}}{n_{..}}\right)$ (8.9.1)

This is a multiplicative model where μ_{ij} is obtained as the product of the mean cell count and the two ratios $\dfrac{In_{i.}}{n_{..}} = \dfrac{n_{i.}}{(n_{..}/I)}$ and $\dfrac{Jn_{.j}}{n_{..}} = \dfrac{n_{.j}}{(n_{..}/J)}$.

The first ratio is the observed row count $(n_{i.})$ as a proportion of the expected row count $\left(\dfrac{n_{..}}{I}\right)$ under the Null Model. The second ratio is the observed column count $(n_{.j})$ as a proportion of the expected column count $\left(\dfrac{n_{..}}{J}\right)$ under the Null Model.

Taking logarithms of both sides of (8.9.1) gives an equation of the form:

$$\log(\mu_{ij}) = \lambda + \lambda_{Ri} + \lambda_{Cj} \qquad (8.9.2)$$

where λ represents the logarithm of the mean cell count,

λ_{Ri} is a term derived from the row proportion $\pi_{i.}$

and λ_{Cj} is a term derived from the column proportion $\pi_{.j}$.

Equation (8.9.2) may be extended to give a Saturated Model which contains a parameter for every cell in the contingency table and produces a model which fits the data perfectly i.e. $\mu_{ij} = n_{ij}$ *for all i, j* .

The Saturated Model (the model containing all the possible terms) is:

$$\log(\mu_{ij}) = \lambda + \lambda_{Ri} + \lambda_{Cj} + \lambda_{RCij} \qquad (8.9.3)$$

where the λ_{RCij} are association terms that reflect deviations from independence of rows and columns. This is a log-linear model i.e. the logarithm of the dependent variable is expressed as a linear function of the other parameters.

The log-linear model (8.9.3) may be extended to contingency tables with more than two dimensions. For a five dimensional table with factors A, B, C, D, E the saturated model would be of the form:

$$\begin{aligned}
\log(\mu_{ijklm}) = {} & \lambda && \text{(1 constant term)} \\
& + \lambda_{Ai} + \lambda_{Bj} + \lambda_{Ck} + \lambda_{Dl} + \lambda_{Em} && \text{(5 first order terms)} \\
& + \lambda_{ABij} + \lambda_{ACik} + \dots\dots\dots + \lambda_{DElm} && \text{(10 second order terms)} \\
& + \lambda_{ABCijk} + \lambda_{ABDijl} + \dots\dots\dots + \lambda_{CDEklm} && \text{(10 third order terms)} \\
& + \lambda_{ABCDijkl} + \lambda_{ABCEijkm} + \dots\dots\dots + \lambda_{BCDEjklm} && \text{(5 fourth order terms)} \\
& + \lambda_{ABCDEijklm} && \text{(1 fifth order term)}
\end{aligned}$$

For the present study with five factors: Yield Group (Y), Bank (B), Lot Size (L), Province (P) and Time (Early or Late Years) (T), and with i, j, k, l and m representing the levels of the corresponding factors, the saturated model can be written as follows.

$$\begin{aligned}
\log(\mu_{ijklm}) = \quad & \lambda & \text{(1 constant term)} \\
& + \lambda_{Yi} + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} & \text{(5 first order terms)} \\
& + \lambda_{YBij} + \lambda_{YLik} + \dots\dots\dots + \lambda_{PTlm} & \text{(10 second order terms)} \\
& + \lambda_{YBLijk} + \lambda_{YBPijl} + \dots\dots\dots + \lambda_{LPTklm} & \text{(10 third order terms)} \\
& + \lambda_{YBLPijkl} + \lambda_{YBLTijkm} + \dots\dots\dots + \lambda_{BLPTjklm} & \text{(5 fourth order terms)} \\
& + \lambda_{YBLPTijklm} & \text{(1 fifth order term)}
\end{aligned}$$

8.9.3 Introduction to analysis

8.9.3.1 Model selection

As shown in the previous section, a log-linear model of the *Provinces A or B* data may include up to thirty-one terms in the predictor variables. Each of these terms may be included or excluded from a particular model. The number of possible models is therefore 2^{31} = 2,147,843,648

However, only sixteen of the terms in the model include the dependent variable *Yield Group*. Therefore, the number of possible models containing *Yield Group* is 2^{16} = 65,536.

It is not cost effective to evaluate all possible models to determine which is best for current purposes. It is therefore important to find a method of selecting as simple as possible a model, which remains adequate for giving good predictions. The method chosen was to use backwards elimination. This involved starting with the saturated model, which is the full model, and sequentially removing the least influential terms until all the remaining terms were assessed as being valuable. The values of the terms were assessed according to the Akaike Information Criterion (AIC) (Akaike, 1974).

The Akaike Information Criterion (AIC) was developed and subsequently presented by Hirotsugu Akaike in 1974 as a way of establishing the goodness of fit of an estimated model (Akaike, 1974). Instead of being a tool for hypothesis testing, the AIC is a measure between models. Hence, it is a way of selecting the most appropriate model. It works on the basis of ranking models, based upon the same dataset, according to their AIC. The model shown to have the lowest AIC should be selected as the best model for that particular dataset. The AIC ranks each model according to how closely its fitted values coincide with the actual values.

Since interest centred on the relationship of *Yield Group* with the other variables, backwards elimination was performed only on the terms including *Yield Group*, and all terms that did not include *Yield Group* were retained in the model.

8.9.3.2 Practical and statistical significance

Agresti (1996: 161-162) states that statistically significant effects might not be practically significant. The *Three Main Banks B* and the *Three Main Banks C* datasets are large samples. With large samples, an effect that may be statistically significant might not necessarily be practically significant (Agresti). According to Agresti, it is therefore important to concentrate on estimation instead of hypothesis testing. Hence, it is important to focus on the accuracy of the model predictions rather than exclusively on which terms are included in the model.

8.9.4 Analysis of the *Provinces A* dataset

8.9.4.1 Log-linear regression output generated from the Saturated Model for *Provinces A*

The output from the log-linear analysis with respect to the saturated model for Provinces A produced an Analysis of Deviance table which is reproduced in Table 8.25.

Table 8.25 Analysis of Deviance Table for the Provinces A Saturated Model

Residual	Df	Residual Deviance	Df	Deviance	P(>|Chi|)
Including 5th order term	0	3.1713e-10			
Excluding 5th order term	6	5.8147e-10	-6	-2.6434e-10	1

The Analysis of Deviance table shows that the fifth order term is not at all significant. P-values fall into the range 0 to 1. A low p-value (0.1 or less) would be taken as an indication that the fifth order term is significant. The p-value found is very high (indistinguishable from 1) showing very clearly that the term is not significant.

Having established that the fifth order term was not at all significant, other terms could be removed using backward elimination. All the terms not containing *Yield Group* were retained. Thus the fourth order term *Provinces A*Lot Size Group*Bank*Time* and all the lower order terms in these variables were retained.

Backward elimination, using the Akaike Information Criterion (AIC) was used to remove terms involving *Yield Group*. This resulted in a basic model with ten terms, including *Yield Group*Provinces A*Lot Size Group*Bank*, which is the highest order term that includes *Yield*. This term was tested specifically in order to ascertain how statistically significant it was. On the basis of a chi-square (χ^2) test it was shown to be significant at the one per cent level. So this term was retained in the basic model.

8.9.4.2 Equation for the basic model for ***Provinces A***

The terms involving *Yield Group* in combination with other factors in the basic model for the *Provinces A* dataset were:

$$\begin{aligned} \log(\mu_{ijlkm}) = \; & \lambda \\ & + \lambda_{YBij} + \lambda_{YLik} + \lambda_{YPil} + \lambda_{YTim} \\ & + \lambda_{YBLijk} + \lambda_{YBPijl} + \lambda_{YBTijm} + \lambda_{YLPikl} \\ & + \lambda_{YBLPijkl} \end{aligned} \tag{8.9.4}$$

For this model the predicted counts in the ninety-six cells are very close to the actual counts. The results for this basic model are included in the log-linear output with respect to *Provinces A*, which is reproduced in Appendix VII. The fourth cell in the contingency table had the largest residual difference, having a predicted count of 20.04 for an actual count of 19. Therefore, the largest residual for the model was 1.04.

None of the coefficients, which again are shown in the output reproduced in Appendix VII, is significant, but the model is hugely significant. The basic model (on seventy-four degrees of freedom) reduces the deviance by 1,689.9348. This is shown in Table 8.26.

Table 8.26 Analysis of Deviance table for the basic model for Provinces A

```
    Null deviance: 1696.3925  on 95  degrees of freedom
Residual deviance:    6.4577  on 21  degrees of freedom
AIC: 354.44
```

The difference in deviance constitutes a test statistic which is chi-squared distributed on seventy-four degrees of freedom (Agresti, 1996: 96). This statistic is significant at well beyond the 0.1 per cent level. Statistically, the model is highly significant.

8.9.4.3 Summary of the basic model for *Provinces A*

The largest residual was 1.04. The model was significant at the 0.1% level, but none of the co-efficients was significant. In all cases, the standard errors of the co-efficients were orders of magnitude greater than the estimated co-efficient values.

8.9.4.4 Log-linear regression output generated from the Simplified Model for ***Provinces A***

The fourth order term comprising *Yield Group, Bank, Lot Size Group* and *Provinces A* in equation 8.9.4 was retained in the basic model as being statistically significant. That term was then tested for practical significance. In order to do this, the fourth order term was removed from the model before the re-starting of the backward elimination. This procedure resulted in a simplified model that included five terms which contained *Yield Group* in combination with one of the other factors. The model including these five terms is shown in section 8.9.4.5 below.

8.9.4.5 Equation for the simplified model for ***Provinces A***

The simplified equation for the reduced model found after the second sequence of backward elimination is shown below:

$$\begin{aligned}\log(\mu_{ijklm}) = \; & \lambda \\ & + \lambda_{YBij} + \lambda_{YLik} + \lambda_{YPil} + \lambda_{YTim} \\ & + \lambda_{YBTijm}\end{aligned} \tag{8.9.5}$$

An explicit test of the three least significant terms in yield according to AIC was undertaken to ascertain how statistically significant they were. The term *Yield Group:Time:Bank* was found to be significant at beyond the 0.1% level. The term *Yield Group:Lot Size* was found to be significant at the 1% level. The term *Yield Group:Provinces A* was found to be significant at beyond the 0.1% level.

The Analysis of Deviance table for the simplified model for *Provinces A* is shown in Tables 8.27.

Table 8.27 Analysis of Deviance table for the simplified model for Provinces A

```
    Null deviance: 1696.393  on 95  degrees of freedom
Residual deviance:   33.534  on 38  degrees of freedom
AIC: 347.52
```

This shows that the simplified model for *Provinces A* is statistically significant at well beyond the 0.1 per cent level.

Many of the individual coefficient estimates in this simplified model were highly significant. These coefficients will be commented on in more detail in the section on logistic regression.

The fitted cell frequencies in the simplified model are somewhat different from those that were produced by the basic model. The largest residual for the simplified model is 1.77. This means that the simplified model is not quite as good as the basic model in terms of the largest residual. Nevertheless, the cell frequencies in the simplified model are still generally quite a good fit. Practically, the predicted cell frequencies are quite close to the actual cell frequencies. The prediction for cells with large counts is generally better than for those cells with small counts. The prediction for large cells is usually within 1 of the actual count. The predicted count almost always differs from the actual count by less than one (<1). So, from a practical viewpoint, the simplified model looks almost as good as the basic model for *Provinces A*.

8.9.4.6 Summary of the simplified model for *Provinces A*

The model was highly significant on a chi-square (x^2) test. Many of the individual co-efficient values were highly significant. The maximum residual was 1.77. For cells with large counts, the predicted count almost always differed from the actual count by <1.

8.9.4.7 Conclusions for *Provinces A* Log-linear regression

The basic model has nine terms which include *Yield Group* in combination with other factors. *Yield Group* is related to the other variables through nine terms, which are:

- Yield Group*Bank
- Yield Group*Time
- Yield Group*Lot Size Group
- Yield Group*Provinces A
- Yield Group*Bank*Time
- Yield Group*Bank*Lot Size Group
- Yield Group*Bank*Provinces A
- Yield Group*Lot Size Group*Provinces A
- Yield Group*Bank*Lot Size Group*Provinces A

The implication is that *Yield Group* is influenced by the other predictor variables in these nine terms i.e. *Bank, Lot Size Group, Provinces A* and *Time*.

The simplified model only has five terms which include *Yield Group* in combination with the other factors. The five terms linking *Yield Group* to other variables are:

- Yield Group*Bank
- Yield Group*Time
- Yield Group*Lot Size Group
- Yield Group*Provinces A
- Yield Group*Time*Bank

This implies that *Yield Group* is influenced by the other variables in these five terms. The relevant influences are the four linear effects plus the one second order effect, the *Time*Bank* interaction.

An examination of the actual cell counts compared with the predicted cell counts from the basic and the simplified models suggests that the simplified model may in practical terms be just as useful as the more complex basic model.

8.9.5 Analysis of the *Provinces B* dataset

8.9.5.1 Log-linear regression output generated from the Saturated Model for *Provinces B*

The output from the log-linear analysis with respect to the saturated model for Provinces B produced an Analysis of Deviance table which is reproduced in Table 8.28

Table 8.28 Analysis of Deviance Table for the Provinces B Saturated Model

Residual	Df	Residual Deviance	Df	Deviance	P(>|Chi|)
Including 5th order term	0	2.1674e-10			
Excluding 5th order term	6	3.0204e-10	6	-8.5301e-11	1

The Analysis of Deviance table shows that the fifth order term is not at all significant.

Having established that the fifth order term was not at all significant, other terms could be removed using backward elimination. All the terms not containing *Yield Group* were retained. Thus the fourth order term *Provinces B*Lot Size Group*Bank*Time* and all the lower order terms in these variables were retained.

Backward elimination, using the Akaike Information Criterion (AIC) was used to remove terms involving *Yield Group*. This resulted in a basic model with thirteen terms, including *Yield Group*Provinces B*Lot Size Group*Bank* and *Yield Group*Provinces B*Lot Size Group*Time*, which were the highest order terms that included *Yield*.

The term *Yield Group*Provinces B*Lot Size Group*Bank* was tested specifically in order to ascertain how statistically significant it was. On the basis of a chi-square (χ^2) test it was shown to be significant at the one per cent level. So this term was retained in the basic model.

The term *Yield Group*Provinces B*Lot Size Group*Time* was tested specifically in order to ascertain how statistically significant it was. On the basis of a chi-square (χ^2) test it was shown not to be significant at the ten per cent level. So this term was removed from the basic model.

8.9.5.2 Equation for the basic model for *Provinces B*

The terms involving *Yield Group* in combination with other factors in the basic model for the *Provinces B* dataset were:

$$\begin{aligned}\log(\mu_{\text{ijklm}}) \quad = \quad & \lambda \\ & + \lambda_{YBij} + \lambda_{YLik} + \lambda_{YPil} + \lambda_{YTim} \\ & + \lambda_{YBLijk} + \lambda_{YBPijl} + \lambda_{YBTijm} + \lambda_{YLPikl} + \lambda_{YLTikm} + \lambda_{YPTilm} \\ & + \lambda_{YBLPijkl}\end{aligned} \tag{8.9.6}$$

This basic model comprised twelve terms, including the constant term.

For this model the predicted counts in the ninety-six cells are very close to the actual counts. The results for this basic model are included in the log-linear output with respect to *Provinces B*, which is reproduced in Appendix VIII. The largest residual difference in the contingency table was 0.96.

Only four of the coefficients, which again are shown in the output reproduced in Appendix VIII, were significant, but the model is hugely significant. The basic model (on seventy-eight degrees of freedom) reduces the deviance by 1,517.6205. This is shown in Table 8.29

Table 8.29 Analysis of Deviance table for the basic model for Provinces B

```
    Null deviance: 1523.6595  on 95  degrees of freedom
Residual deviance:    6.0386  on 17  degrees of freedom
AIC: 373.44
```

The difference in deviance constitutes a test statistic which is chi-squared distributed on seventy-four degrees of freedom (Agresti, 1996: 96). This statistic is significant at well beyond the 0.1 per cent level. Statistically, the model is highly significant.

8.9.5.3 Summary of the basic model for *Provinces B*

The largest residual was 0.96. The model was significant at the 0.1% level, but only four of the model co-efficients was significant at the ten per cent level. In all other cases, the standard errors of the co-efficients were orders of magnitude greater than the estimated co-efficient values.

8.9.5.4 Log-linear regression output generated from the Simplified Model for *Provinces B*

The fourth order term *Yield Group, Bank, Lot Size Group* and *Provinces B* in equation 8.9.6 was retained in the basic model as being statistically significant. That term was then tested for practical significance. In order to do this, the fourth order term was removed from the model before the re-starting of the backward elimination. This procedure resulted in a simplified model that included five terms which contained *Yield Group* in combination with one of the other factors. The model including these five terms is shown in section 8.9.5.6 below.

8.9.5.5 Equation for the simplified model for ***Provinces B***

$$
\begin{aligned}
\log(\mu_{ijklm}) \quad = \quad & \lambda \\
& + \lambda_{YBij} + \lambda_{YLik} + \lambda_{YPil} + \lambda_{YTim} \\
& + \lambda_{YBTijm}
\end{aligned}
\qquad (8.9.7)
$$

An explicit test of the three least significant terms in yield according to AIC was undertaken to ascertain how statistically significant they were. The term *Yield Group:Time:Bank* was found to be significant at beyond the 0.1% level. The term *Yield Group:Lot Size* was found to be significant at the 5% level. The term *Yield Group:Provinces B* was found to be significant at beyond the 0.1% level.

The Analysis of Deviance table for the simplified model for *Provinces B* is shown in Table 8.30

Table 8.30 Analysis of Deviance table for the simplified model for Provinces B

```
    Null deviance: 1523.659  on 95  degrees of freedom
Residual deviance:   35.304  on 38  degrees of freedom
AIC: 360.71
```

This shows that the simplified model for *Provinces B* is statistically significant at well beyond the 0.1 per cent level.

Many of the individual coefficient estimates in this simplified model were highly significant. These coefficients will be commented on in more detail in the section on logistic regression.

The fitted cell frequencies in the simplified model are somewhat different from those that were produced by the basic model. The largest residual for the simplified model is 2.62. This means that the simplified model is not quite as good as the basic model in terms of the largest residual. Nevertheless, the cell frequencies in the simplified model are still generally quite a good fit. Practically, the predicted cell frequencies are quite close to the actual cell frequencies. The prediction for cells with large counts is generally better than for those cells with small counts. The prediction for large cells is usually within 2 of the actual count. The predicted count almost always differs from the actual count by less than two (<2). So, from a practical viewpoint, the simplified model may be almost as good as the basic model for *Provinces B*.

8.9.5.6 Summary of the simplified model for *Provinces B*

The model was highly significant on a chi-square (x^2) test. Many of the individual co-efficient values were highly significant. The maximum residual was 2.62. For cells with large counts, the predicted count almost always differed from the actual count by <2.

8.9.5.7 Conclusions for *Provinces B* Log-linear regression

The basic model has nine terms which include *Yield Group* in combination with other factors. *Yield Group* is related to the other variables through nine terms, which are:

- Yield Group*Bank
- Yield Group*Time
- Yield Group*Lot Size Group
- Yield Group*Provinces B
- Yield Group*Bank*Time
- Yield Group*Bank*Lot Size Group
- Yield Group*Bank*Provinces B
- Yield Group*Time*Lot Size Group
- Yield Group*Time*Provinces B

- Yield Group*Lot Size Group*Provinces B
- Yield Group*Bank*Lot Size Group*Provinces B

The simplified model only has five terms which include *Yield Group*. The five terms linking *Yield Group* to other variables are:

- Yield Group*Bank
- Yield Group*Time
- Yield Group*Lot Size Group
- Yield Group*Provinces B
- Yield Group*Time*Bank

This implies that *Yield Group* is influenced by the other variables in these five terms. The relevant influences are the four linear effects plus the one second order effect, the *Time*Bank* interaction.

An examination of the actual cell counts compared with the predicted cell counts from the basic and the simplified models suggests that the simplified model may in practical terms be just as useful as the more complex basic model.

8.9.6 Discussion of the log-linear analysis

In log-linear analysis, the dependent variable, *Yield*, has been transformed into the categorical variable *Yield Group*. During the log-linear analysis, *Yield Group* is not actually treated as a dependent variable. Instead, the log-linear analysis simply identifies meaningful relationships between the variables which may give pointers to which combinations of variables influence the *Yield* value.

Only those banking-halls let to the three main banks were subjected to the log-linear analysis in order to reduce the number and the proportion of empty cells. In so doing, it was possible to retain seven-ninths of the data whilst reducing the number of cells in the contingency table to just ninety-six.

Basic and simplified models were built with respect to the datasets relating to both *Provinces A* and *Provinces B*. These were built using backward elimination from the respective saturated models using the Akaike Information Criterion. The respective two simplified models were built after removing those terms that had been in the basic models as statistically significant, but subsequently suspected not to be practically significant.

The basic model with respect to *Provinces B* is slightly more complex than that for *Provinces A*, having two additional interaction terms. However, both the *Provinces A* and the *Provinces B* datasets produced exactly the same simplified model. These two simplified models suggested that Yield is influenced by the four first order terms and by the *Bank*Time* interaction. The influencing factors are shown in the equation.

The findings of the literature review, the theoretical perspectives section and the qualitative study confirmed the findings of the log-linear analysis. The log-linear analysis, in turn, confirmed those independent variables being useful predictors of *Yield*.

The log-linear analysis strongly suggests that there is a linear relationship between *Yield Group* and *Bank*, *Time*, *Lot Size Group* and *Provinces*. Furthermore, in the respective datasets containing *Provinces A* and *Provinces B*, *Yield* is influenced by the interaction between *Bank* and *Time*. Also, the suggestion is that the simplified model generated from the *Three Main Banks B* and the *Three Main Banks C* datasets may be as practically useful as the more complex basic models for the two datasets.

8.10 Final conclusions from the Exploratory Data Analysis

The final conclusions arising from the Exploratory Data Analysis can be summarized as follows:

1. Examination of the *Purified Freehold* data showed that the dataset lacked balance. The number of premises sold varied markedly across *Banks, Regions, Lot Size* and *Years*.
2. Attempted cross-tabulation of the *Purified Freehold* dataset showed that the data were very sparse. In order to reduce the sparseness, the data were collapsed to reduce the number of levels for each of the predictor factors.

3. In section 7.3 *Yield* was defined as a numeric variable. Following the data collapse which produced the *Three Main Banks* datasets, *Yield* was transformed into a factor, *Yield Group*, with two levels derived from the bi-modal yield distribution.
4. As part of the data collapse process, the geographical variable *Region* was transformed into a provinces variable which took two forms. These were *Provinces A* and *Provinces B*. Further analysis is required to establish whether either of these breakdowns is to be preferred for predictive purposes.
5. In the *Three Main Banks B* and *Three Main Banks C* datasets, all four predictor variables (*Bank, Province, Lot Size* and *Time*) are categorical.
6. A chi-squared test for independence of factors in the cross-tabulation contingency table produced a highly significant result indicating that the factors were not statistically independent. Thus, at least one of the factors was influenced by one or more of the other factors. There was clearly a need for further investigation to determine whether yield was being influenced by one or more of *Bank, Province, Lot Size* or *Time* singularly or in combination.
7. The lineplots of Mean Yield against *Time* for *Lot Size* and *Province* suggest that yield may vary across lot sizes and between provinces.
8. The lineplot of Mean Yield against *Time* for the *Three Main Banks* shows no clear indication that yield varies between banks.
9. The lineplot of the number of bank sales against *Time* for the *Three Main Banks* indicates very clearly that the data lack balance with HSBC having sold most of its premises early in the study period, whilst Barclays sold most of its premises late in the study period.
10. The boxplots strongly suggested that *Yield* varied across *Banks, Provinces, Lot Size* and *Time*. The lack of balance in the data made it difficult to draw any firm conclusions about the causes of these variations.
11. Tests of normality showed that there were significant departures from normality in the yield distribution in some cells of the cross-tabulation contingency table. Further analysis is required to investigate the utility of variance stabilizing transformations.
12. The tests of homogeneity of variance showed that the variance of the yield distribution was not constant across cells in the cross-tabulation contingency table.
13. Log-linear analyses of the *Three Main Banks B* and the *Three Main Banks C* datasets suggested that *Yield* is influenced by the levels of the factors *Bank, Province, Lot Size* and

Time and by the effect of the interaction between *Bank* and *Time*. Other terms have a statistically significant effect on *Yield*, but may not have a practical effect.

14. Log-linear analysis identified which predictor variables influenced *Yield*, but did not provide a specific prediction of *Yield* for each case in the dataset. Logistic regression using the same input data would provide predictions, and it was concluded that logistic regression analyses of the *Three Main Banks B* and *C* datasets should be carried out in order to obtain these predictions. In log-linear and logistic regression, all the variables are categorical.
15. With *Yield* as a numeric variable, prediction of the yield values can be obtained using Analysis of Variance (ANOVA) with the predictor variables being categorical. The departure from normality in the data and, in particular, the lack of homogenous variance would mean that the results of the ANOVA might be subject to serious errors, since ANOVA is based on the assumption of normality and constant variance in the data analyzed. It was concluded, however, that an ANOVA should be carried out to establish if it produced results which had practical value for forecasting.
16. The sparseness of the *Three Main Banks* datasets means that there is insufficient data for separate test and validation datasets. Hence, cross-validation must be used.
17. Also, it is clear that over time, the predictor variable, *Year*, has an effect on another predictor, *Lot Size*. The reason for this is two-fold, Firstly, there was a trend for the yields on all United Kingdom investment property to diminish over time during the period. Conversely, there was a trend for the rents of premises to rise during the same period. Either one of these factors will increase capital values and hence *Lot Size*. These two factors combined in a shift from the number of lots that fell into the lower lot sizes to the number of lots into the higher lot sizes over the course of the period subject to investigation.

9 Advanced Analyses I: Original dataset incorporating calendar data

9.1 Introduction

The purpose of this research is to create a predictive framework which investors can use to maximise investment yields in retail bank premises. There are two bases of models in the framework as follows:

1. The main intention is to predict *Yield Group*, which is what the majority of investors want to know. Logistic regression is a statistical procedure that would enable this.

2. If an ANOVA model can be developed, it would allow investors to go a step further in being able to assess the yields and then the value at which they should bid by capitalizing rents at the appropriate yield.

The original data made available by the auction houses about British banking-halls sold at auction as investments, provided a date variable. Due to the sparseness and lack of balance in the data, the Exploratory Data Analysis chapter found that the data needed to be collapsed sufficiently to render it capable of meaningful analysis. As part of that collapse, calendar date was collapsed as a *Time* factor, represented by early and late years. This chapter therefore examines logistic regression and ANOVA models using *Time* as an independent variable. Subsequent analysis substituting the Investment Property Databank (IPD) index for time is examined in a further chapter.

9.2 Logistic Regression

9.2.1 Introduction to logistic regression

According to Agresti (1996: 162-167), there is a one-to-one equivalence between log-linear models and binary logistic regression models fitted to the same data. Accordingly, the *Three Main Banks* datasets from the log-linear analysis were used for the logistic regression. The relationship between the variable, *Yield Group*, and the other variables revealed by the two models are identical. However, a difference exists inasmuch as although log-linear shows where relationships exist, no variable is taken as a dependent variable (Agresti, 1996: 145). With logistic regression, the relationship is expressed in a more useable form in that:

- *Yield Group* appears explicitly as a dependent variable; and

- The fitted logistic equation estimates the probability that each case in the data set belongs to the high yield group.

As a result of *Yield Group* becoming a dependent variable, it appears explicitly on the left-hand side of the equation as a value to be predicted from the other variables. The results from binary logistic regression therefore provide useful predictions for making investment decisions.

Since the Exploratory Data Analysis had shown that Province was a significant variable whether *Region* was collapsed into *Provinces A* or into *Provinces B*, both forms of collapsed data need to be modelled and tested. Therefore, an adequate model needed to be found separately with respect to both *Provinces A* and *Provinces B*. Accordingly, two sets of modelling and testing were undertaken.

As when fitting the best model for log-linear, it is necessary to fit a best model for logistic regression. The use of backward elimination allows firstly the removal of terms not having a statistical significance and subsequently those terms having a statistical but not a practical significance (Agresti, 1996: 162). Thus, it allows the fitting of a less complex model. It allows the fitting of a model that is adequate using the least number of necessary terms. The Akaike Information Criterion (AIC) was the procedure used to remove terms using backward elimination.

According to Barlas (1996), a study's external validity relies upon the validity of any model used. He shows that the validity of a model is based upon the validity of its purpose. Barlas makes the case against formal model validation prior to analysis in socio-economic research that is entirely focused on empirical data. This is because the model is considered to be valid if its output corresponds with reality within the conventions of degrees of statistical significance for the discipline in question (Barlas). Barlas states that models of prediction, including forms of regression, fall into this type. The purpose of the models built in the logistic regression used in this study is to predict high or low *Yield Group*. Accordingly, a case is made for checking the validity of the models built after analyses through cross-validation. Validity is about whether the model's output is a true reflection of the purpose of that model. That is to say that to have model validity, a model needs to correlate well with the real world situation. The cross-validation that was undertaken in the analyses checked the models' correlation with reality. Predictive accuracy is different. Predictive accuracy is about how well the model predicts in different circumstances, which in the case of this study is under different economic conditions reflected by *Time*. Hence, the difference is between descriptive validity and predictive accuracy.

Normally, validation is undertaken by taking a further sample that is distinct from the data used in the model-building. The norm for such validation is for there to be a like split between the dataset used in the model-building and the sample of further data used in the validation. However, the current study relies on datasets that comprise all the available data for the study period. That means that validation had to be undertaken using samples taken from the totality of data used in the model-building. The model-building was performed using the entire population. Such testing of the model through the use of other samples is known as cross-validation (Field, 2005:171). Field argues that for a model to remain valid, it must be able to generalize and therefore retain accuracy in its predictive power across different samples.

Data splitting, relying on the random division of the data into two equal parts and the comparison of the predictive power of the resultant models is an ideal form of cross-validation (Field, 2005:172). However, there is seldom sufficient data to enable such an approach (Field, 2005: 172). Statistical techniques have been developed to overcome such lack of data. One such statistical technique, known as *bootstrapping*, was introduced by Bradley Efron (Kohavi, 1995). Bootstrapping relies upon an iterative process of testing a randomly split sample of a dataset against the rest of that dataset (Efron and Gong (1983) and Kohavi (1995)). Bootstrapping is accepted as a means of cross-validation (Efron and Tibshirani, 1986). Hoesli *et al.* (1997: 3) show how bootstrapping has previously been used to test validity of datasets relating to property investment markets. Bootstrapping is a form of cross-validation that randomly uses uniform sets of samples split without replication (Kohavi, 1995; Efron and Tibshirani, 1997).

The leave-one-out model selection cross-validation approach can lead to less reliable estimates for linear models due to high variance (Kohavi, 1995). Ten-fold and five-fold cross-validation have been shown to have a lower variance than the leave-one-out cross-validation and, therefore, to provide a higher degree of confidence (Efron and Tibshirani, 1997). Accordingly, cross-validation was undertaken using ten lots of 9:1 samples randomly taken without replication. The original dataset was randomly divided into ten sub-sets without replication with the omission of one of the ten sub-sets at a time while the remaining nine-tenths of the original dataset were used to estimate the parameters of the model. This gave a prediction of membership of the high yield group independent of cases whose probability was being estimated. The predictive accuracy of the cross-validation of each model is reproduced in the respective confusion matrices in section 9.2.4.

9.2.2 The Model for the *Provinces A* and *Provinces B* datasets

The full, or saturated, binary logistic model for the two datasets with dependent variable *Yield Group* (Y) and the four factors, *Bank* (B), *Lotsize* (L), *Province* (P) and *Time* (T) as predictor variables takes the form:

$$
\begin{aligned}
\log\left(\frac{p_i}{1-p_i}\right) = \; & \lambda && \text{(1 constant term)} \\
& +\lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} && \text{(4 first order terms)} \\
& +\lambda_{BLjk} + \lambda_{BPjl} + \dots\dots\dots + \lambda_{PTlm} && \text{(6 second order terms)} \\
& +\lambda_{BLPjkl} + \lambda_{BLTjlm} + \dots\dots\dots + \lambda_{LPTklm} && \text{(4 third order terms)} \\
& +\lambda_{BLPTjklm} && \text{(1 fourth order term)}
\end{aligned}
\qquad (9.1)
$$

where p_i is the predicted probability that the i'th case belongs to the high yield group, and where *j, k, l* and *m* represent the categories of the factors *B, L, P* and *T* respectively.

Equation 9.1 can be re-arranged to give:

$$p_i = \exp(f(\lambda)) / (1 + \exp(f(\lambda)))$$

where

$$\begin{aligned} f(\lambda) \quad = \quad & \lambda \\ & + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\ & + \lambda_{BLjk} + \lambda_{BPjl} + \dots\dots\dots + \lambda_{PTlm} \\ & + \lambda_{BLPjkl} + \lambda_{BLTjlm} + \dots\dots\dots + \lambda_{LPTklm} \\ & + \lambda_{BLPTjklm} \end{aligned} \qquad (9.2)$$

9.2.3 Analysis

9.2.3.1 Analysis of Provinces A

Backward elimination from the saturated model gave a basic model:

$$\begin{aligned} f(\lambda) \quad = \quad & \lambda \\ & + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\ & + \lambda_{BLjk} + \lambda_{BPjl} + \lambda_{BTjm} + \lambda_{LPkl} \\ & + \lambda_{BLPjkl} \end{aligned} \qquad (9.3)$$

Cross-validation was then performed on the basic model in order to ascertain how good its results were in practical terms.

The statistical significance of the third order term, *Bank*Lot Size*Provinces A*, was then examined. The term was found to be statistically significant. In order to ascertain if it was practically significant, it was removed prior to the continuation of the backward elimination. That produced a simplified model:

$$\begin{aligned} f(\lambda) \quad = \quad & \lambda \\ & + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\ & + \lambda_{BTjm} \end{aligned} \qquad (9.4)$$

The significance of the *Lot Size* term, which was the least significant term in the model, was tested for statistical significance. It was significant at the 1.0 per cent level, so it was retained in the model. A cross-validation of the simplified model was then undertaken. The results for the basic and simplified models were then compared.

9.2.3.2 Analysis of Provinces B

Backwards elimination from the saturated model produced a basic model comprising twelve terms in the predictor variables. The significance of the *Bank*Lot Size*Provinces B* term was tested. It was significant at the 1.0 per cent level, so it was retained in the model. The *Time*Lot Size*Provinces B* term was then tested. This latter term was not significant at the 10 per cent level, so it was removed from the model. The resultant model was:

$$\begin{aligned} f(\lambda) = {} & \lambda \\ & + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\ & + \lambda_{BLjk} + \lambda_{BPjl} + \lambda_{BTjm} + \lambda_{LPkl} + \lambda_{LTkm} + \lambda_{PTlm} \\ & + \lambda_{BLPjkl} \qquad (9.5) \end{aligned}$$

Cross-validation was then performed on the basic model.

Then in order to test the practical significance of the remaining third order term, it was removed from the model and the backward elimination continued. This produced a simplified model comprising five terms in the predictor variables. The equation for the simplified model was:

$$\begin{aligned} f(\lambda) = {} & \lambda \\ & + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\ & + \lambda_{BTjm} \qquad (9.6) \end{aligned}$$

The statistical significance of the three least significant terms in the simplified model was tested. The *Lot Size* term was significant at the 5 per cent level and was, therefore, retained.

The *Provinces B* term was significant at the 0.1 per cent level and was, therefore, retained. The *Bank*Time* interaction term was significant at the 0.1 per cent level and so was retained.

Cross-validation of the simplified model was undertaken. Comparisons were then made between the cross-validation results for the basic and simplified models.

9.2.3.3 Comparison of *Provinces A* and *Provinces B*

After completing the analyses for *Provinces A* and *Provinces B*, the results for the two forms of province collapse were compared.

9.2.4 Results

9.2.4.1 Provinces A

Backwards elimination from the saturated model produced a basic model (equation 9.3) comprising nine terms in the predictor variables:

$$
\begin{aligned}
f(\lambda) = \; & \lambda \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\
& + \lambda_{BLjk} + \lambda_{BPjl} + \lambda_{BTjm} + \lambda_{LPkl} \\
& + \lambda_{BLPjkl}
\end{aligned}
$$

The output from the logistic regression with respect to the basic model for Provinces A produced an Analysis of Deviance table which is reproduced in Table 9.1.

Table 9.1 Analysis of Deviance Table for the Provinces A Basic Model

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	
NULL			690	953.86		
Bank	2	253.745	688	700.12	< 2.2e-16	***
Time	1	211.606	687	488.51	< 2.2e-16	***
Lot size	1	3.285	686	485.22	0.069906	.
Provinces A	3	39.545	683	445.68	1.330e-08	***
Bank:Time	2	28.771	681	416.91	5.657e-07	***
Bank:Lot size	2	1.079	679	415.83	0.582953	
Bank:Provinces A	6	5.201	673	410.63	0.518348	
Lot Size:Provinces A	3	0.774	670	409.85	0.855669	
Bank:Lot size:Provinces A	5	20.022	665	389.83	0.001238	**

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

The linear terms *Bank*, *Time*, and *Provinces A* were all found to be significant at well beyond the 0.1% level. The linear term *Lot Size* was shown in the Pearson chi-square within the Analysis of Deviance table to be significant at the 10% level. The second order term *bank:time* was also found to be significant at well beyond the 0.1% level. The three second order terms *Bank:Lot Size, Bank:Provinces A* and *Lot Size:Provinces A*, although useful according to the Aikake Criterion, were shown not to be significant at the 10% level. The single third order term *Bank:Lot Size:Provinces A* was shown to be significant at the 1% level.

The coefficients table, reproduced in the *Provinces A* logistic regression output in Appendix VII, showed that none of the coefficients was at all significant. This is surprising given that the model itself is shown by the Reduced Analysis of Variance table (Table 9.2) to be quite a good fit. The deviance table shows that the model is significant at well beyond the 0.1 per cent level (chi-squared is 564.03 on 35 degrees of freedom). Hence, it would be very dangerous to interpret what is happening on the basis of these coefficients. They cannot be used for any confident interpretation.

Table 9.2 Reduced Analysis of Deviance for the basic model for Provinces A

```
    Null deviance: 953.86  on 690  degrees of freedom
Residual deviance: 389.83  on 665  degrees of freedom
AIC: 441.83
```

Null deviance is the residual produced after the null model, which only comprises a constant term, has been fitted. The alternative model is that being fitted.

Confusion matrices were produced to summarize the results in a compact, easily understood way. A confusion matrix is a table that shows predicted classifications against actual classifications. Since the present analysis builds a model for predicting those banking-hall lots that fall into the high *Yield Group,* there is a need to identify the model where the predicted and observed high *Yield Groups* most nearly coincide. The best model is therefore the one which produces confusion matrices where the predicted and actual classifications coincide the most.

A confusion matrix plots the actual membership of the high and low *Yield Groups* against the predicted membership of those groups for a given cut-off probability or threshold. From the confusion matrix it is possible to derive the true positive rate and the false positive rate with the given threshold. Hence, the plot will place each case into one of four boxes in the matrix. Those corrected predicted to be in the high and low *Yield Groups* will be placed into the *True Positive* and the *True Negative* boxes respectively. Those wrongly predicted will be placed in either the *False Positive* or the *False Negative* boxes.

The basic model produced the following confusion matrix:

Table 9.3 Confusion Matrix of Provinces A Basic Model: p = 0.5

		Actual	
		Low Yield	High Yield
Predicted	Low Yield	288	12
	High Yield	84	307

The original sample, *Three Main Banks B,* was divided into ten random sub-sets without replication, and the analysis then dropped out one of the ten at a time to be the test sample whilst the other nine-tenths are used to estimate the parameters of the model. The procedure gave a prediction of high yield group membership based on a model whose derivation was independent of cases whose probability was being estimated. The procedure is called cross-validation.

It is important that the randomness of bootstrapping is borne in mind. Since bootstrapping splits the sample into ten sub-sets randomly every time for cross-validation, the way in which the sample is split can be expected to be different every time. Due to the split being different each time that the test is undertaken, marginally different results may result on each occasion. This is quite normal and in no way reduces the validity of the cross-validation.

After cross-validation, the basic model for *Provinces A* produces a probability bar chart which is reproduced in Figure 9.1.

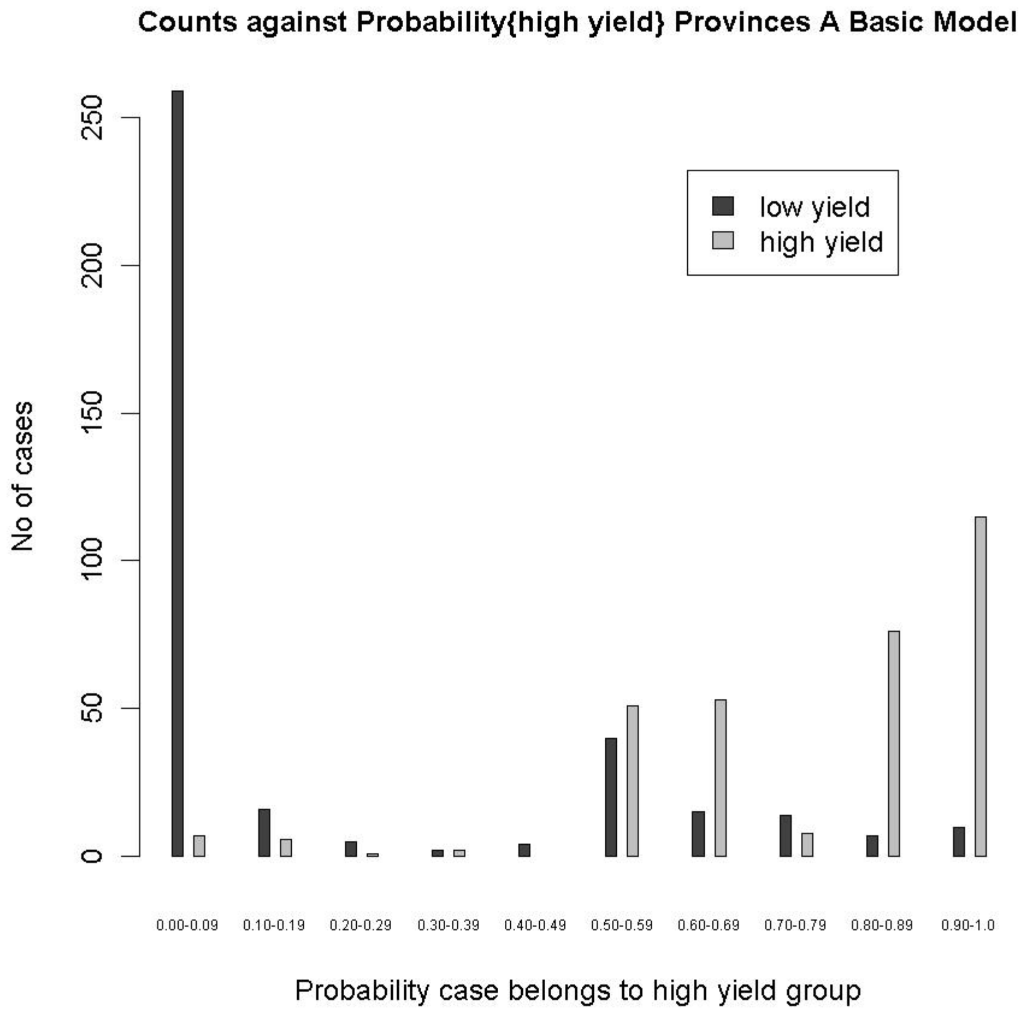


Figure 9.1 Probability bar chart for the basic model for Provinces A after cross-validation

This gives rise to the following confusion matrix in Table 9.4 which is very similar to Table 9.3.

Table 9.4 Confusion Matrix of Provinces A Basic Model after Cross-validation: p = 0.5

		Actual	
		Low Yield	High Yield
Predicted	Low Yield	286	16
	High Yield	86	303

ROC (Receiver Operating Characteristics) curves are commonly used to compare the predictive accuracy of alternative models. They lend themselves to the checking of the accuracy of binary classification. To create a ROC curve, the true positive rate is plotted as a function of the false positive rate at different possible decision thresholds (Fawcett, 2006).

True positive rate = true positive/(true positive + false negative).

False positive rate = false positive/(true negative + false positive).

In the current analysis, the high *Yield Group* is treated as the positive one so that the ROC chart can plot the proportion of positives correctly predicted.

Following the cross-validation of the basic model with respect to *Provinces* A, a ROC chart was produced to. It is capable of being used to identify the best threshold for any model. This ROC chart is shown in Figure 9.2.

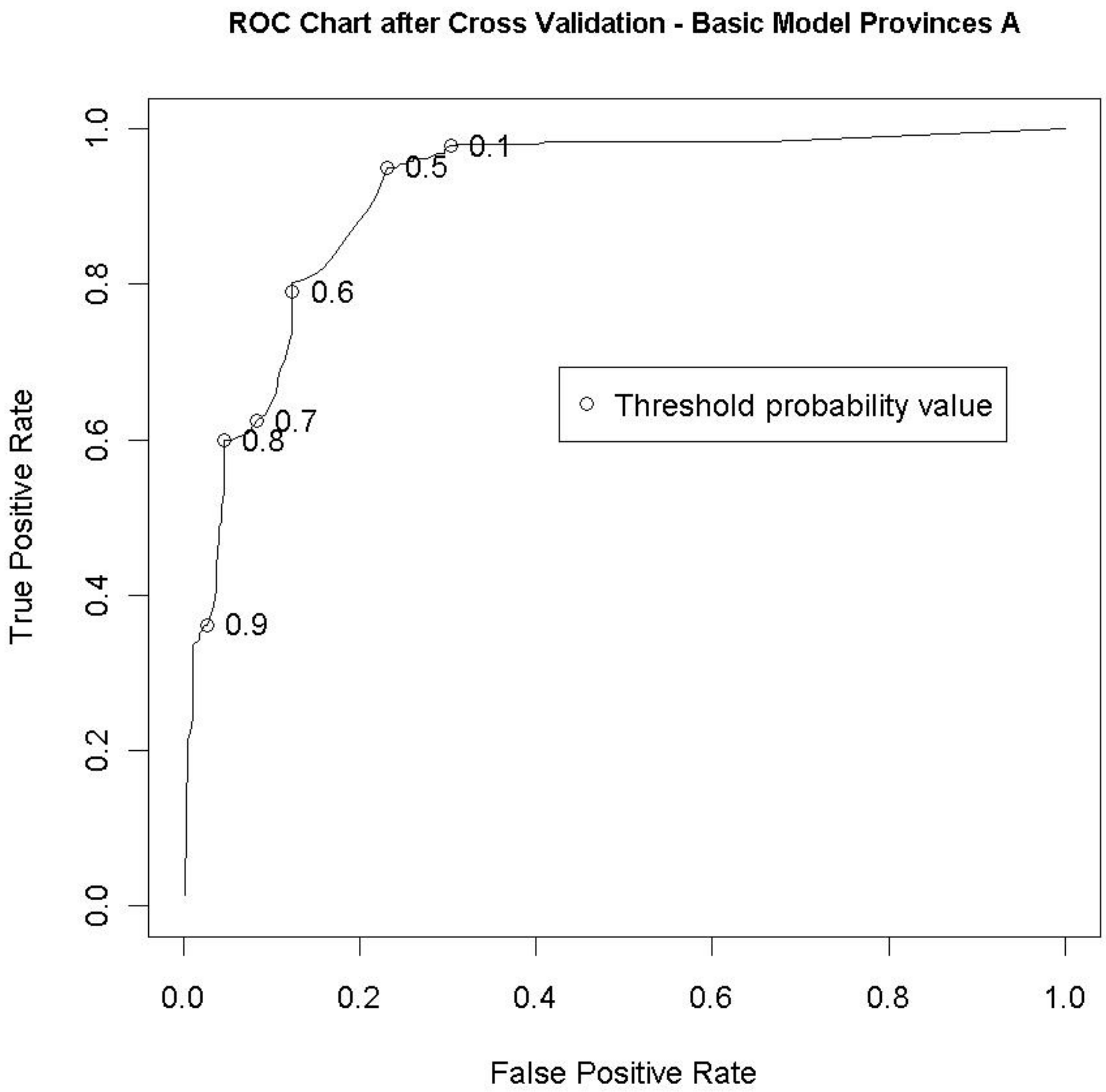


Figure 9.2 ROC Chart after Cross-Validation for the Provinces A Basic Model

Examination of the predictions bar chart and the ROC chart suggest that for an acceptably low (0.1 or less) false positive rate, a threshold of 0.8 or more is required.

To test whether the third order term was practically significant, it was removed from the basic equation and the backward elimination process re-started. This produced a simplified equation (equation 9.7), containing five terms in the predictor variables.

$$f(\lambda) = \lambda + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} + \lambda_{BTjm} \tag{9.7}$$

The output from the logistic regression with respect to the simplified model for Provinces A produced an Analysis of Deviance table which is reproduced in Table 9.5.

Table 9.5 Analysis of Deviance Table for the Provinces A Simplified Model

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	
NULL			690	953.86		
Bank	2	253.745	688	700.12	< 2.2e-16	***
Time	1	211.606	687	488.51	< 2.2e-16	***
Lot size	1	3.285	686	485.22	0.069906	.
Provinces A	3	39.545	683	445.68	1.330e-08	***
Bank:Time	2	28.771	681	416.91	5.657e-07	***

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

The Reduced Analysis of Deviance table is given in Table 9.6.

Table 9.6 Reduced Analysis of Deviance for the simplified model for Provinces A

```
    Null deviance: 953.86  on 690  degrees of freedom
Residual deviance: 416.91  on 681  degrees of freedom
AIC: 436.91
```

A chi-squared test (536.95 on 9 degrees of freedom) shows the model to be significant at well beyond the 0.1 per cent level.

The linear terms *Bank*, *Time*, and *Provinces A* were all found to be significant at the 0.1% level (see Table 9.7). The second order term *bank:time* was also found to be significant at the 0.1% level.

The linear term *Lot size* was shown in the Pearson chi-square within the Analysis of Deviance table (Table 9.5) to fall outside the five percent significance value at 6.99 per cent. This is because the significances are dependent on the order in which the terms are inserted and the true significance of a term is shown by the removal of that term from the model. Therefore, the significance of *Lot Size* was tested explicitly (see Table 9.7), whereupon it was found to be significant at the 1% level at 0.8853 per cent.

Table 9.7 Analysis of Deviance Table – Lot Size Provinces A

```
 Resid.              Df Resid. Dev     Df  Deviance
P(>|Chi|)
Simplified Model       681   416.91
Simplified - LotSize    682    423.76  -1  -6.8523
0.008853
```

Table 9.8 Coefficients Table for the Provinces A Simplified Model

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	15.7819	700.1381	0.023	0.982016	
Bank HSBC	-12.0676	700.1381	-0.017	0.986248	
Bank LTSB	-14.0908	700.1380	-0.020	0.983943	
Time Late	-20.9332	700.1387	-0.030	0.976148	
Lot Size Large	0.7759	0.3067	2.530	0.011396	*
Provinces A: Midlands	-0.8786	0.4446	-1.976	0.048145	*
Provinces A: South	-1.5738	0.4144	-3.797	0.000146	***
Provinces A: Celt	0.9820	0.6159	1.594	0.110835	
Bank HSBC: Time Late	13.9879	700.1392	0.020	0.984060	
Bank LTSB: Time Late	17.8488	700.1388	0.025	0.979661	

Significance codes 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1

The coefficients table, re-produced in Table 9.8, produced nine entries, there being fewer coefficients fitted than in the basic model. Some of the individual coefficients in this simplified model are significant. The differences between the two banking companies, HSBC and Lloyds TSB, and between these two banks and Barclays, are not significant at the 10 per cent level. The difference between the early and late time coefficients is not shown to be significant, being well outside the 10% level. The differences between the *Bank:Time* interaction coefficients are also shown not be significant at the ten per cent level. Thus, although the Bank, Time and Bank:Time interaction effects are all highly significant (see Table 9.5), none of the individual coefficients is significant (see Table 9.8).

The difference between the *Lot Size* coefficients is shown to be significant at the 5% level. Thus there is a distinction between small and large *Lot Size*.

As far as provinces are concerned, there is a clear distinction between the south and the Midlands on the one hand and the north of England on the other. Furthermore, the effect is shown to be much more significant with the south than it is with the Midlands. With respect to the south of England, the significance is shown to be well beyond the 0.1% level. As far as the Midlands is concerned, the significance only lies beyond the 5.0% level. However, the significance between the north of England and those regions grouped together as the Celt province is not significant at the ten per cent level. Differences amongst the coefficients for South, Midlands and Celt provinces are not significant at the 10% level.

The simplified model produced the following confusion matrix:

Table 9.9 Confusion Matrix of Provinces A Simplified Model: p = 0.5

		Actual	
		Low Yield	High Yield
Predicted	Low Yield	287	12
	High Yield	85	307

After cross-validation the following confusion matrix was produced:

Table 9.10 Confusion Matrix of Provinces A Simplified Model after Cross-validation: p = 0.5

		Actual	
		Low Yield	High Yield
Predicted	Low Yield	287	12
	High Yield	85	307

The two confusion matrices are identical.

After cross-validation, the simplified model for *Provinces A* produced a probability bar chart which is reproduced in Figure 9.3.

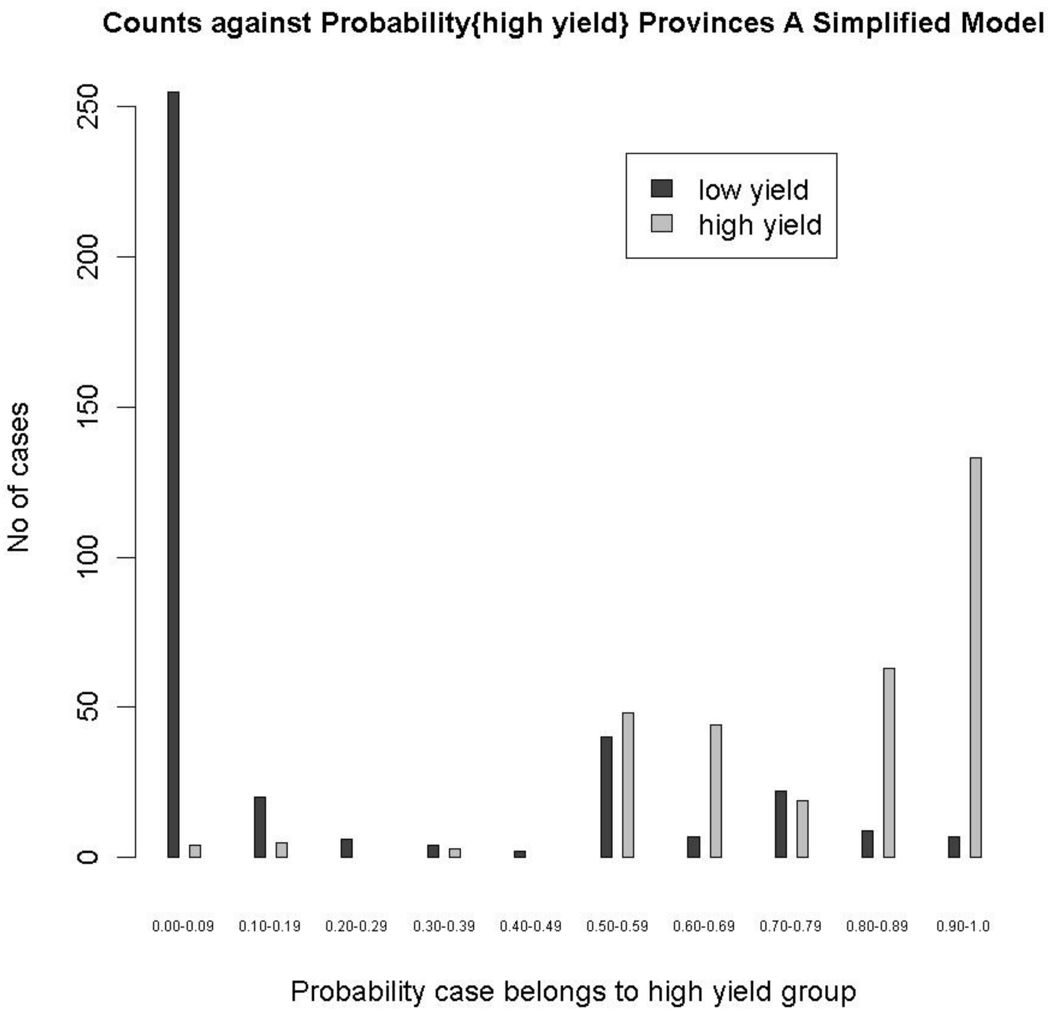


Figure 9.3 Probability bar chart for the simplified model for Provinces A after cross-validation

Following the cross-validation of the simplified model with respect to *Provinces* A, a ROC chart was produced. It is shown in Figure 9.4. This particular ROC chart suggests that 0.8 looks a good threshold based upon a false positive rate not exceeding 0.1.

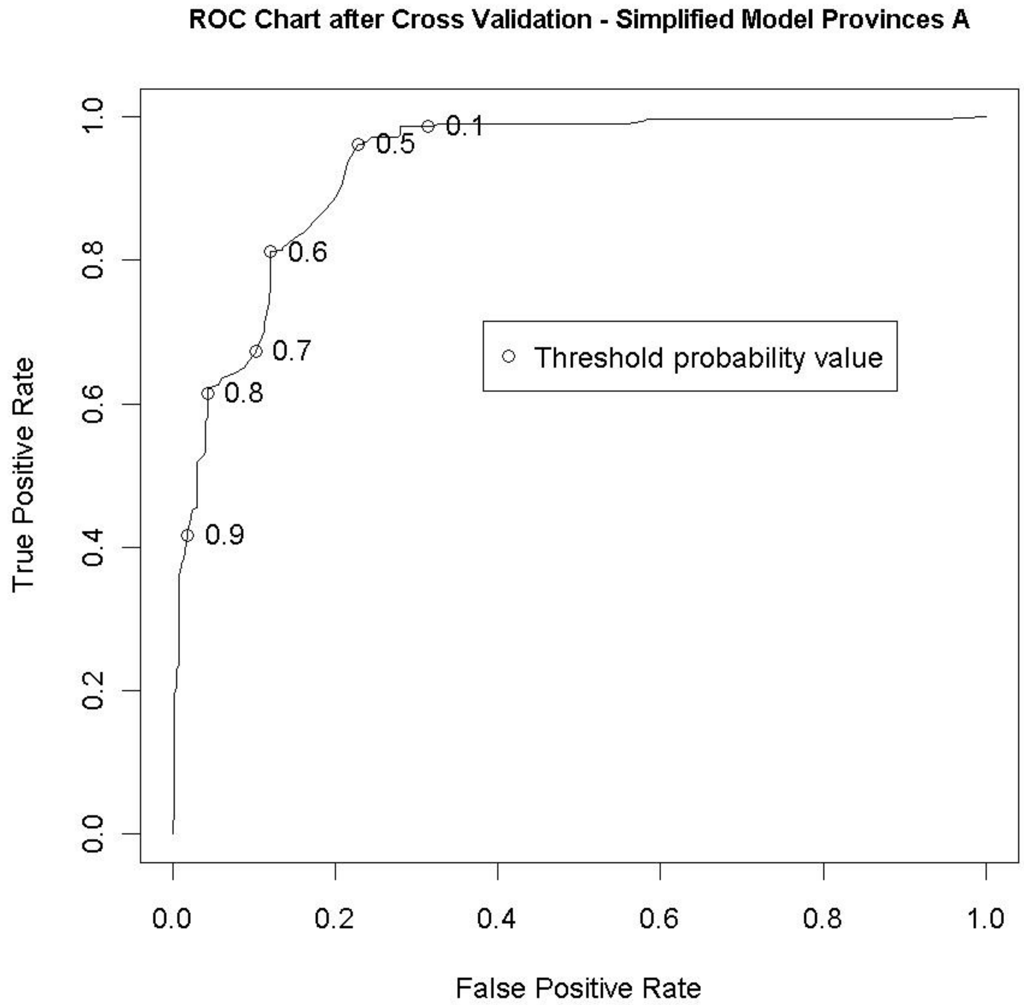


Figure 9.4 **ROC Chart after Cross-Validation for the Provinces A Simplified Model**

To achieve a good decision threshold, it is necessary to satisfy the following criteria:

1. A threshold equal to or greater than 6.35 per cent, since the histogram of *Yield* by the *Three Main Banks* was bimodal in distribution and had shown the high yield group as being equal to or above 6.35 per cent.

2. A predictive positive rate of ninety per cent or more is considered appropriate, because no more than ten per cent of the predictions for high yield should be wrong. This is equivalent to a high true positive/false positive ratio as represented by the Positive Predictive Value (PPV) where:

$$ppv = tp/(tp + fp)$$

where tp = true positive,
fp = false positive

The selection of the predictive positive rate is a subjective judgment decision based upon the risk that a particular investor is prepared to take in portfolio investment selections. For illustrative purposes, in the current case, a value of ninety per cent was considered to be appropriate.

3. The number of actual high yield cases that are predicted as high yield should be as large as possible. This equates with a high true positive rate, which is defined as:

$$tpr = tp/(tp + fn)$$

where tp = true positive,
fn = false negative.

Figure 9.5 shows the Predicted Positive Value versus True Positive Rate after cross-validation with respect to the simplified model for Provinces A.

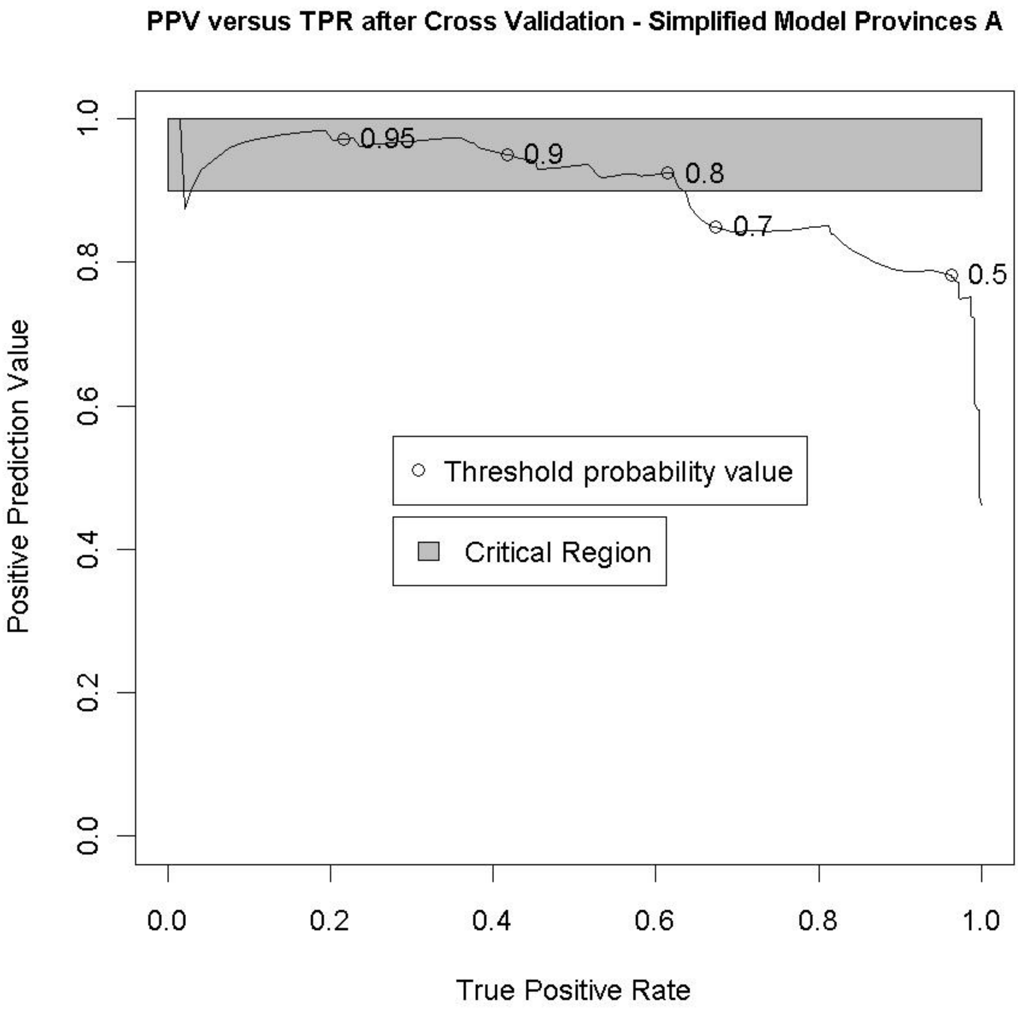


Figure 9.5 **PPV versus TPR after Cross-Validation for the Provinces A Simplified Model**

The erratic behaviour of the graph in Figure 9.5 when the threshold is at an extremely high level very close to the maximum actual or predicted probability, results from the very few actual or predicted cases which are above the threshold. The chart confirms, again, that 0.8 appears to be a good threshold.

Following the generation of the two ROC charts with respect to the basic and simplified models for *Provinces A*, a further ROC chart was produced, comparing the predictive accuracy of the two models. This third ROC chart is shown in Figure 9.6. Clearly, this ROC chart suggests that there is very little difference between the predictive powers of the basic and simplified models with respect to Provinces A. Indeed, an eyeball examination of the chart suggests that the simplified model may actually have marginally greater predictive power than the basic model. The simplified model is, therefore, preferred.

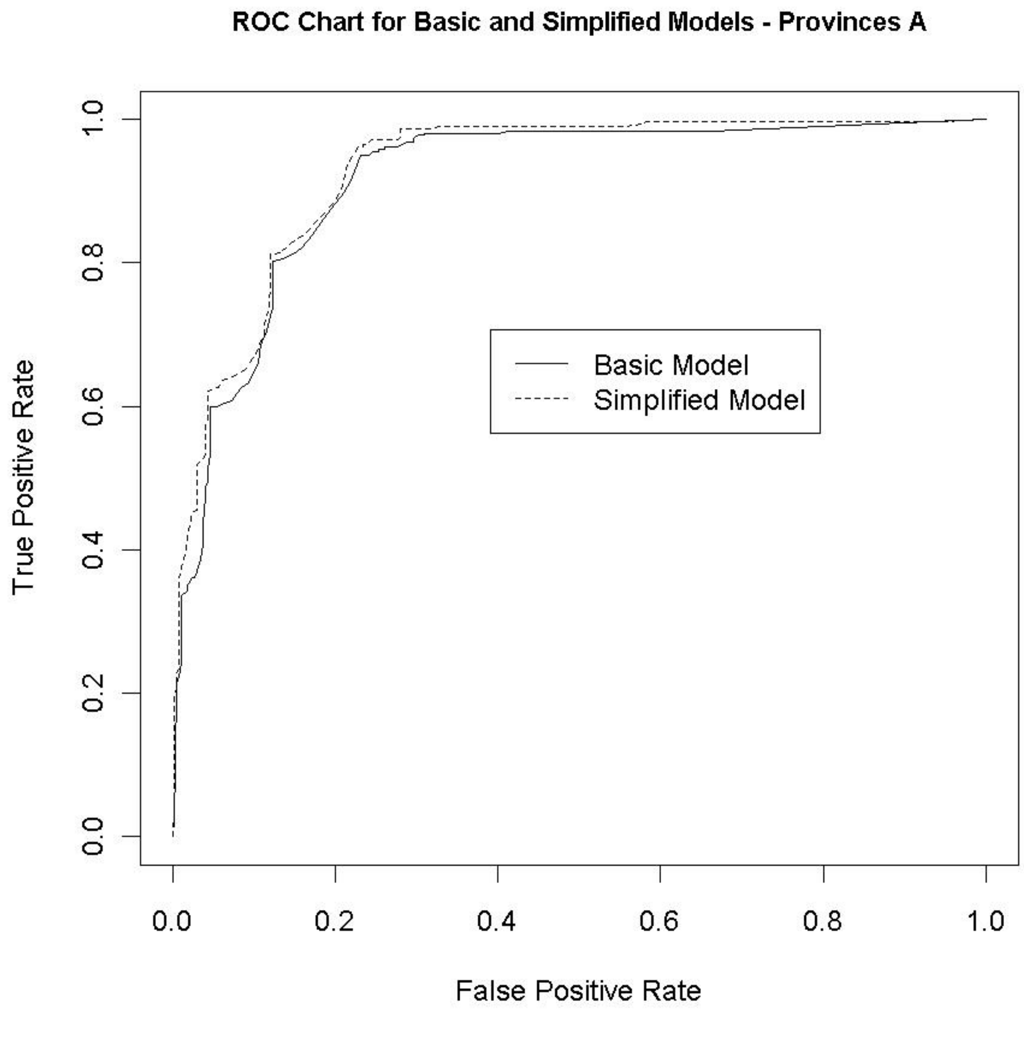


Figure 9.6 **ROC Chart for Provinces A Basic and Simplified Models**

9.2.4.2 Provinces B

Backwards elimination from the saturated model produced a basic model originally comprising twelve terms in the predictor variables. Examination showed that one of the third order terms was not statistically significant. It was therefore removed. This resulted in a model (equation 9.8) containing eleven terms in the predictor variables.

$$\begin{aligned} f(\lambda) = \; & \lambda \\ & + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\ & + \lambda_{BLjk} + \lambda_{BPjl} + \lambda_{BTjm} + \lambda_{LPkl} + \lambda_{LTkm} + \lambda_{PTlm} \\ & + \lambda_{BLPjkl} \end{aligned} \tag{9.8}$$

The output from the logistic regression with respect to the basic model for *Provinces B* produced an Analysis of Deviance table which is reproduced in Table 9.11.

Table 9.11 Analysis of Deviance Table for the Provinces B Basic Model

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	
NULL			690	953.86		
Bank	2	253.745	688	700.12	< 2.2e-16	***
Time	1	211.606	687	488.51	< 2.2e-16	***
Lot size	1	3.285	686	485.22	0.069906	.
Provinces B	3	21.403	683	463.82	8.682e-05	***
Bank:Time	2	26.606	681	437.21	1.669e-06	***
Bank:Lot Size	2	0.861	679	436.35	0.650244	
Bank:Provinces B	6	6.846	673	429.51	0.335345	
Time:Lot size	1	1.012	672	428.50	0.314501	
Time:Provinces B	3	2.774	669	425.72	0.427753	
Lot size:Provinces B	3	2.683	666	423.04	0.443162	
Bank:Lot Size: Provinces B	5	15.090	661	407.95	0.009985	**
Time:Lot size: Provinces B	3	6.039	658	401.91	0.109746	

Significance codes	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’	0.1 ‘ ’	1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

The linear terms *Bank*, *Time*, and *Provinces B* were all found to be significant at well beyond the 0.1% level. The linear term *Lot Size* was shown in the Pearson chi-square within the Analysis of Deviance table to be significant at the 10% level. The second order term *bank:time* was also found to be significant at well beyond the 0.1% level. The four second order terms *Bank:Lot Size, Bank:Provinces B, Time:Lot Size* and *Lot Size:Provinces B*, although useful according to the Aikake Criterion, were shown not to be significant at the 10% level. The third order term *Bank:Lot Size:Provinces B* was shown to be significant at the 1% level. The other third order term *Time:Lot Size:Provinces B* was shown not to be significant at the 10% level. It was, therefore, removed from the basic model.

The coefficients table, reproduced in the *Provinces B* logistic regression output in Appendix VIII, showed that none of the coefficients was at all significant. This is surprising given that the model itself is shown by the Reduced Analysis of Variance table (Table 9.9) to be quite a good fit. The deviance table shows that the model is significant at well beyond the 0.1 per cent level (chi-squared is 545.91 on 29 degrees of freedom). Hence, it would be very dangerous to interpret what is happening on the basis of these coefficients. They cannot be used for any confident interpretation.

Table 9.12 Reduced Analysis of Deviance for the basic model for Provinces B

```
    Null deviance: 953.86  on 690  degrees of freedom
Residual deviance: 407.95  on 661  degrees of freedom
AIC: 467.95
```

Null deviance is the residual produced after the null model, which only comprises a constant term, has been fitted. The alternative model is that being fitted.

Confusion matrices were produced to summarize the results in a compact, easily understood way. The basic model produced the following confusion matrix:

Table 9.13 Confusion Matrix of Provinces B Basic Model p = 0.5

		Actual	
		Low Yield	High Yield
Predicted	Low Yield	284	10
	High Yield	88	309

The original sample, *Three Main Banks C,* was divided into ten random sub-sets without replication, and a cross-validation performed.

After cross-validation, the basic model for *Provinces B* produced a probability bar chart which is reproduced in Figure 9.7.

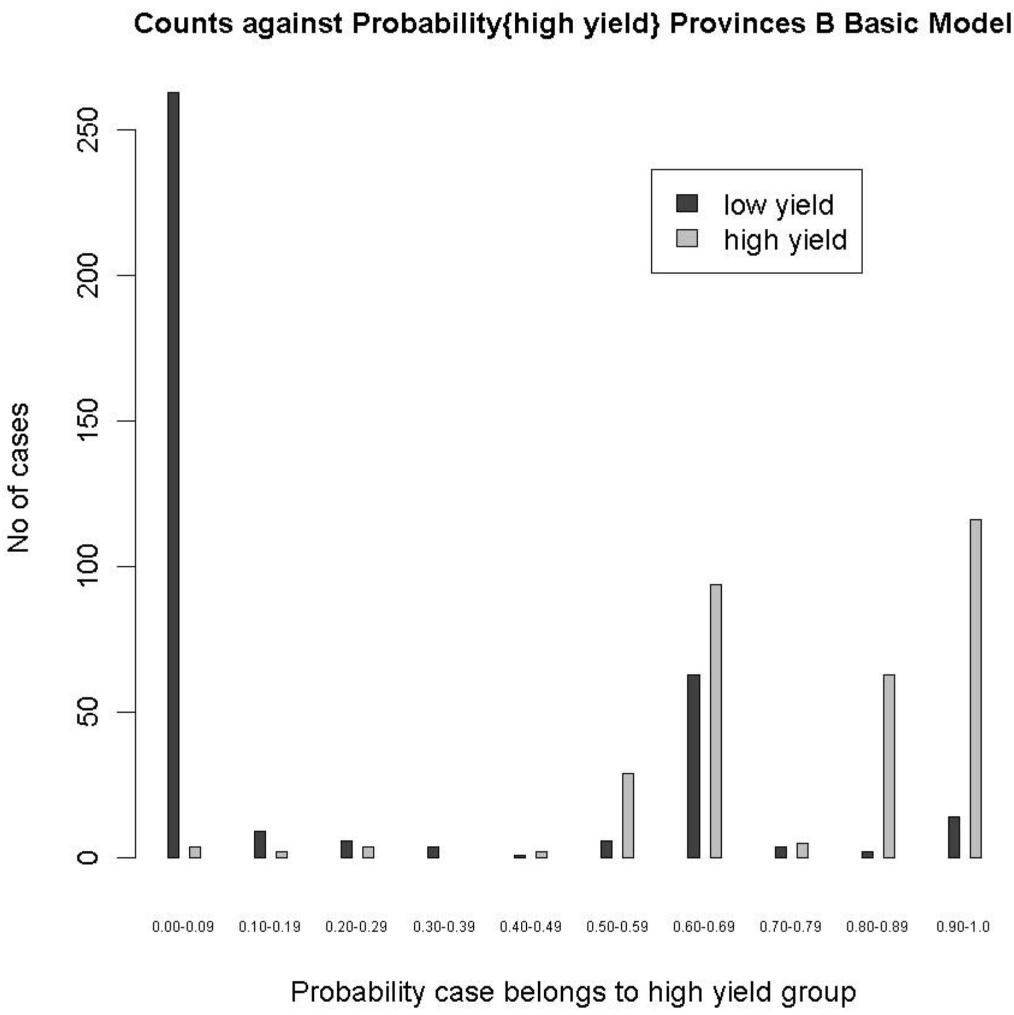


Figure 9.7 Probability bar chart for the basic model for Provinces B

This gives rise to the following confusion matrix in Table 9.14 which is very similar to Table 9.13.

Table 9.14 Confusion Matrix of Provinces B Basic Model after Cross-validation: p = 0.5

		Actual	
		Low Yield	High Yield
Predicted	Low Yield	283	12
	High Yield	89	307

Following the cross-validation of the basic model with respect to *Provinces B*, a ROC chart was produced to. It is capable of being used to identify the best threshold for any model. This ROC chart is shown in Figure 9.8.

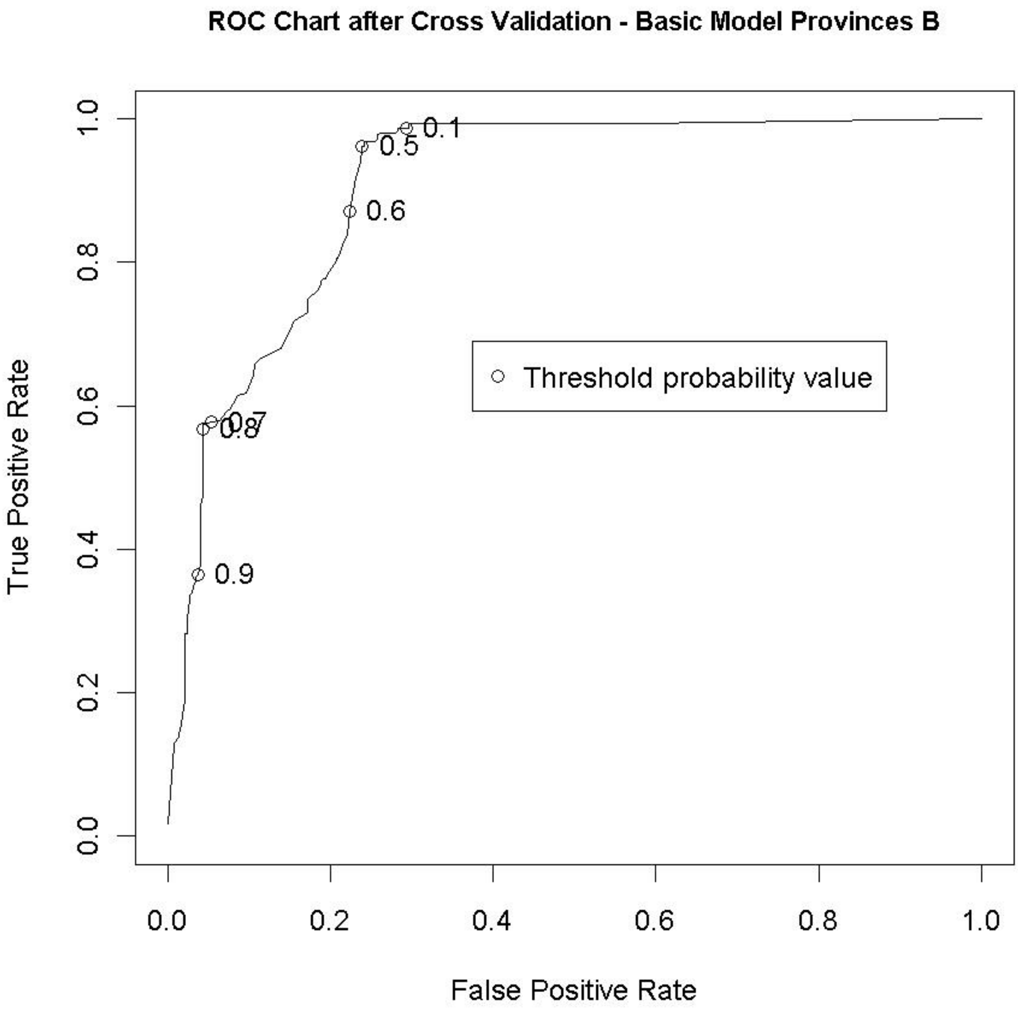


Figure 9.8 ROC Chart after Cross-Validation for the Provinces B Basic Model

Examination of the predictions bar chart and the ROC chart suggest that for an acceptably low (0.1 or less) false positive rate, a threshold of 0.8 or more is required.

To test whether the third order term was practically significant, it was removed from the basic equation and the backward elimination process re-started. This produced a simplified equation (equation 9.9), containing five terms in the predictor variables.

$$
\begin{aligned}
f(\lambda) = {} & \lambda \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\
& + \lambda_{BTjm}
\end{aligned}
\qquad (9.9)
$$

The output from the logistic regression with respect to the simplified model for *Provinces B* produced an Analysis of Deviance table which is reproduced in Table 9.15.

Table 9.15 Analysis of Deviance Table for the *Provinces B* Simplified Model

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	
NULL			690	953.86		
Bank	2	253.745	688	700.12	< 2.2e-16	***
Time	1	211.606	687	488.51	< 2.2e-16	***
Lot size	1	3.285	686	485.22	0.0699	.
Provinces B	3	21.403	683	463.82	8.682e-05	***
Bank:Time	2	26.606	681	437.21	1.669e-06	***
Significance codes	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' '	1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

The Reduced Analysis of Deviance table is given in Table 9.16.

Table 9.16 Reduced Analysis of Deviance for the simplified model for *Provinces B*

```
    Null deviance: 953.86  on 690  degrees of freedom
Residual deviance: 437.21  on 681  degrees of freedom
AIC: 457.21
```

A chi-squared test (516.65 on 9 degrees of freedom) shows the model to be significant at well beyond the 0.1 per cent level.

The linear terms *Bank*, *Time*, and *Provinces B* were all found to be significant at well beyond the 0.1% level (see Table 9.15). The second order term *bank:time* was also found to be significant at well beyond the 0.1% level.

The linear term *Lot size* was shown in the Pearson chi-square within the Analysis of Deviance table (Table 9.15) to fall outside the five per cent significance value at 6.99 per cent. This is because the significances are dependent on the order in which the terms are inserted and the true significance of a term is shown by the removal of that term from the model. Therefore, the significance of *Lot Size* was tested explicitly (see Table 9.17), whereupon it was found to be significant at the 5% level at 3.854 per cent.

Table 9.17 Analysis of Deviance Table – Lot Size *Provinces B*

```
 Resid.                Df Resid. Dev     Df  Deviance
P(>|Chi|)
Simplified Model         681    437.21
Simplified - LotSize      682     441.50  -1  -4.2811
0.03854
```

Table 9.18 Coefficients Table for the *Provinces B* Simplified Model

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	214.2889	697.6201	0.020	0.984	
Bank HSBC	-11.8726	697.6201	-0.017	0.986	
Bank LTSB	-13.9236	697.6200	-0.020	0.984	
Time Late	-20.7362	697.6207	-0.030	0.976	
Lot Size Large	0.6119	0.3038	2.014	0.044	*
Provinces B:Wales &SW	0.1298	0.3094	0.419	0.675	
Provinces B:Midlands	0.4853	0.3276	1.481	0.139	
Provinces B:North Britain	1.7166	0.4232	4.057	4.98e-05	***
Bank HSBC:Time Late	14.0011	697.6212	0.020	0.984	
Bank LTSB:Time Late	17.6507	697.6208	0.025	0.980	

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The coefficients table, re-produced in Table 9.18, produced nine entries, there being fewer coefficients fitted than in the basic model. Some of the individual coefficients in this simplified model are significant. The differences between the two banking companies, HSBC and Lloyds TSB, and between these two banks and Barclays, are not significant at the 10 per cent level. The difference between the early and late time coefficients is not shown to be significant, being well outside the 10% level. The differences between the

Bank:Time interaction coefficients are also shown not be significant at the ten per cent level. Thus, although the Bank, Time and Bank:Time interaction effects are all highly significant (see Table 9.15), none of the individual coefficients is significant (see Table 9.18).

The difference between the *Lot Size* coefficients is shown to be significant at the 5% level. Thus there is a distinction between small and large *Lot Size*.

As far as provinces are concerned, there is a clear significance between North Britain and the other three provinces. This distinction appears to be greatest between North Britain on the one hand and London and the South-east on the other, where the difference is significant at well beyond the 0.1 per cent level. Differences amongst the coefficients for London and the South-east, Wales and the South-west and the Midlands provinces are not significant at the 10% level.

The simplified model produced the following confusion matrix:

Tables 9.19 Confusion Matrix of *Provinces B* Simplified Model: p = 0.5

		Actual	
		Low Yield	High Yield
Predicted	Low Yield	287	13
	High Yield	85	306

After cross-validation the following confusion matrix was produced:

Table 9.20 Confusion Matrix of *Provinces B* Simplified Model after Cross-validation: p = 0.5

		Actual	
		Low Yield	High Yield
Predicted	Low Yield	287	13
	High Yield	85	306

The two confusion matrices are identical.

After cross-validation, the simplified model for *Provinces B* produces a probability bar chart which is reproduced in Figure 9.9.

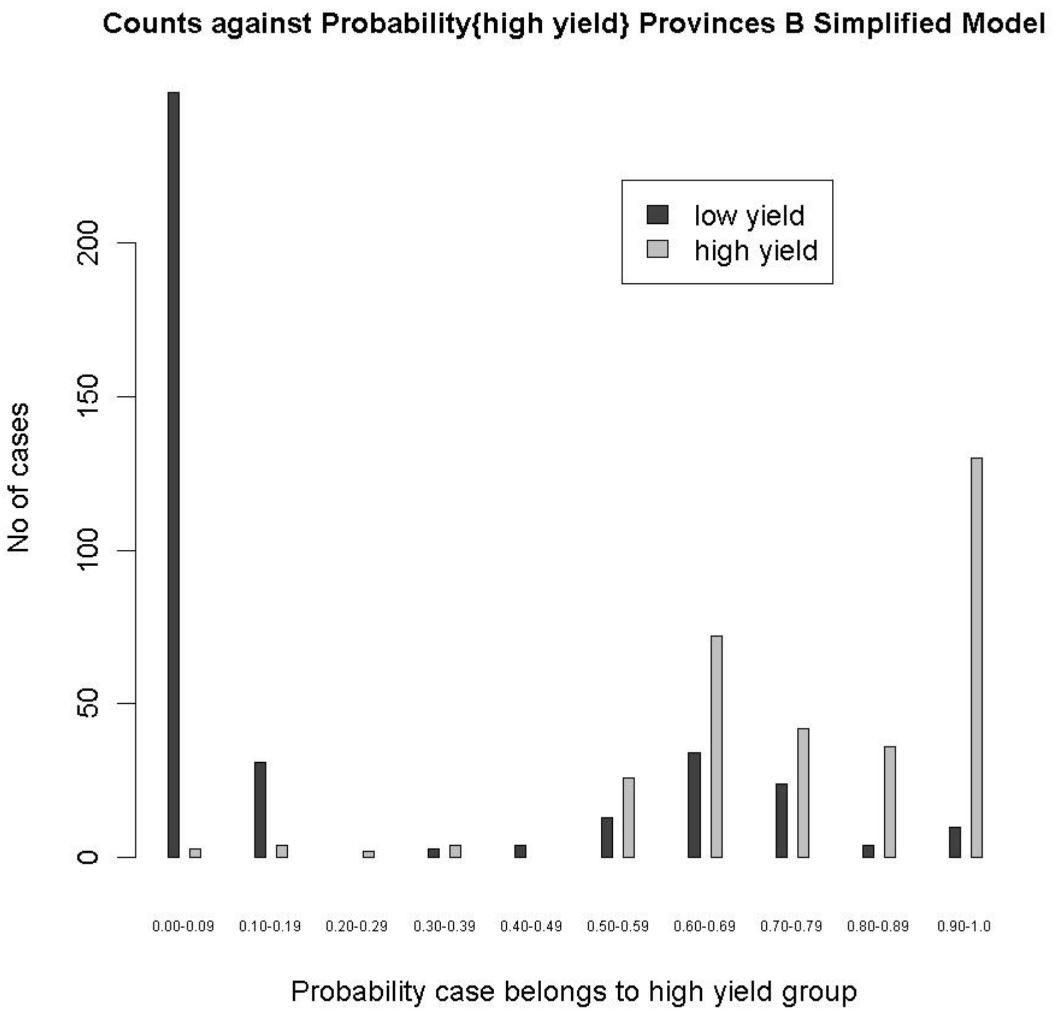


Figure 9.9 Probability bar chart for the simplified model for *Provinces B* after cross-validation

Following the cross-validation of the simplified model with respect to *Provinces B*, a ROC chart was produced. It is shown in Figure 9.10. This particular ROC chart suggests that 0.8 looks a good threshold based upon a false positive rate not exceeding 0.1.

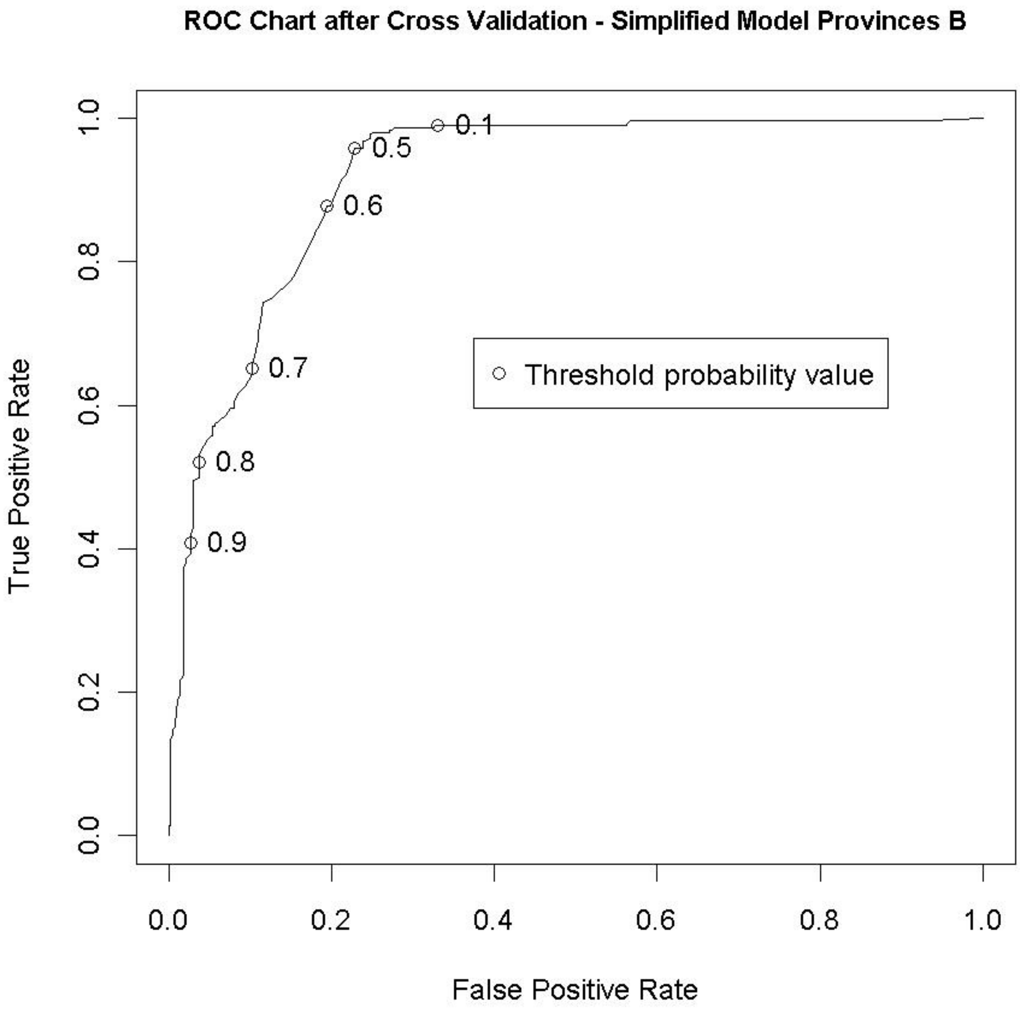
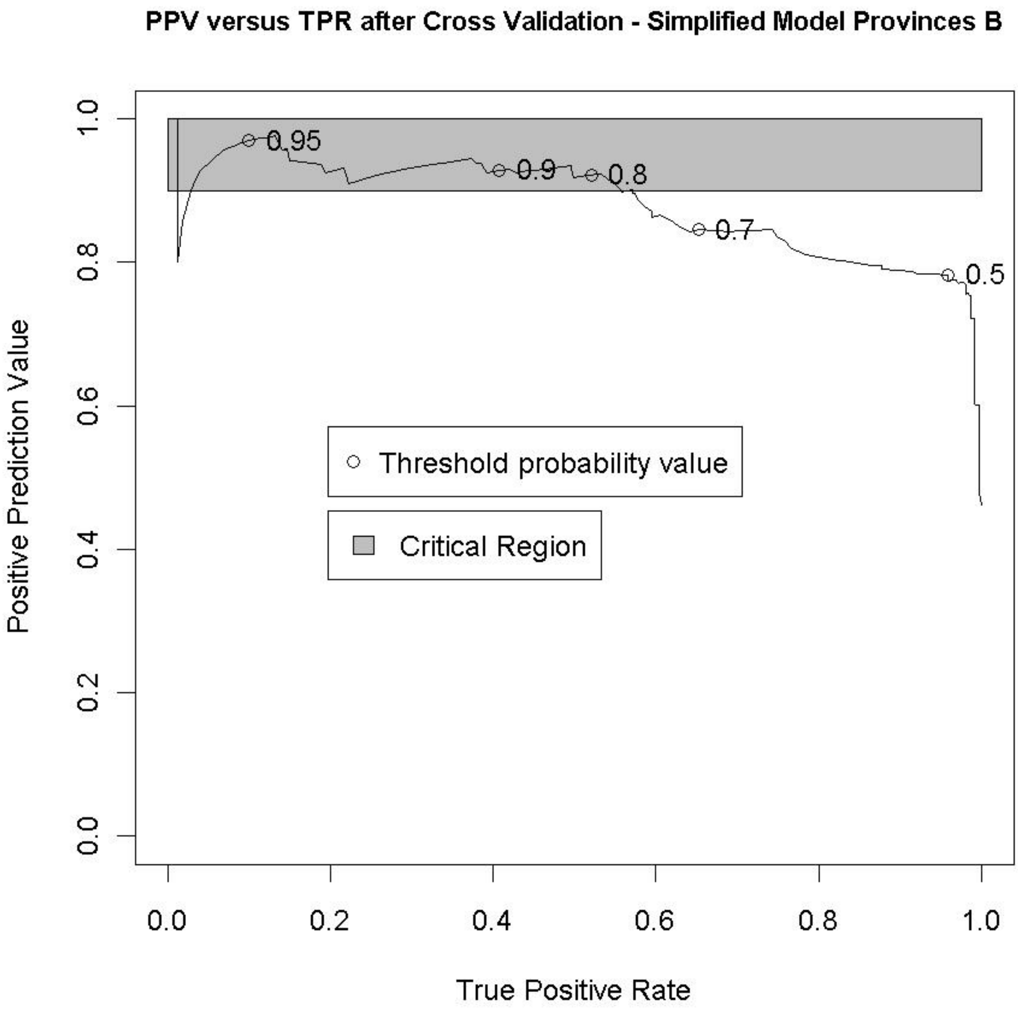


Figure 9.10 ROC Chart after Cross-Validation for the *Provinces B* Simplified Model

Figure 9.11 shows the Predicted Positive Value versus True Positive Rate after cross-validation with respect to the simplified model for Provinces B.



9.11 PPV versus TPR after Cross-Validation for the Provinces B Simplified Model

The erratic behaviour of the graph in Figure 9.11 when the threshold is at an extremely high level very close to the maximum actual or predicted probability results from the very, very few actual or predicted cases which are above the threshold. The chart confirms, again, that 0.8 appears to be a good threshold.

Following the generation of the two ROC charts with respect to the basic and simplified models for *Provinces B*, a further ROC chart was produced, comparing the predictive accuracy of the two models. This third ROC chart is shown in Figure 9.12. Clearly, this ROC chart suggests that there is very little difference between the predictive powers of the basic and simplified models with respect to *Provinces B*. Indeed, an eyeball examination of the chart suggests that the simplified model may actually have marginally greater predictive power than the basic model. The simplified model is, therefore, preferred.

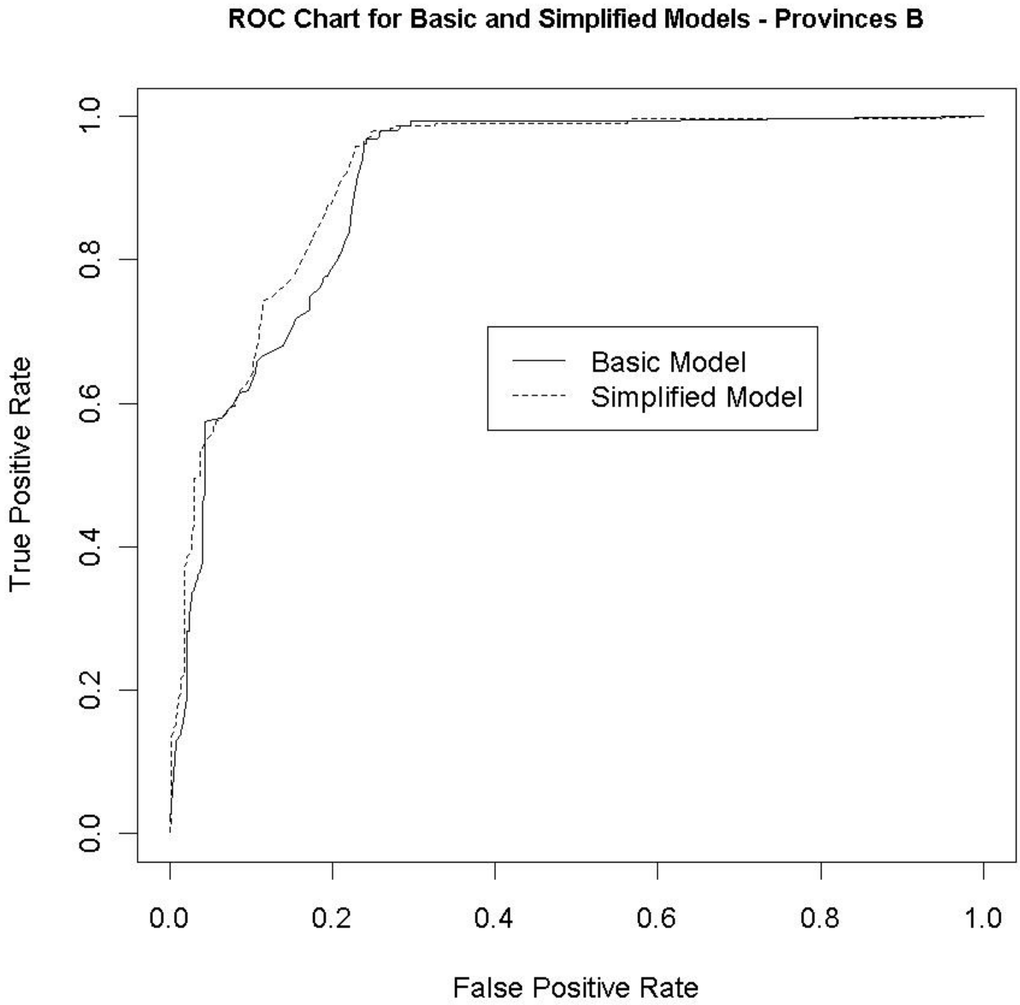


Figure 9.12 ROC Chart for the *Provinces B* Basic and Simplified Models

9.2.4.3 Provinces A and B compared

The ROC graphs for the simplified models of Provinces A and B are reproduced in Figure 9.13.

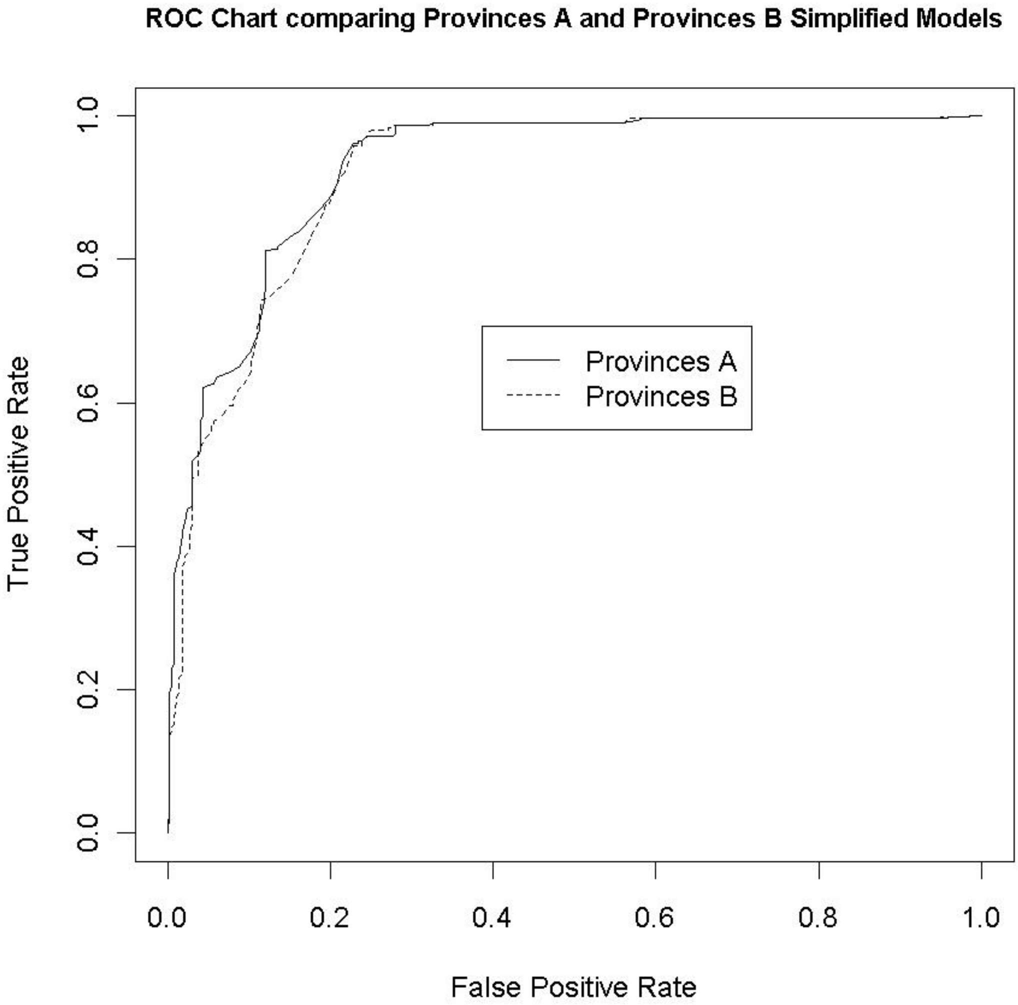


Figure 9.13 ROC Chart comparing *Provinces A* and *Provinces B* Simplified Models

The two curves in the ROC chart in Figure 9.13 are broadly very similar. However, at the region of interest, which is thresholds above 0.5, the curves appear to show that the results for *Provinces A* are marginally better than those for *Provinces B*.

9.2.5 Conclusions

For both datasets, the simplified model was shown to be no worse than the basic model and also comprised less terms. Therefore, the simplified model was preferred on the grounds of simplicity.

The simplified model for both datasets comprised the same terms. They comprised all the linear terms plus the *Bank*Time* interaction.

In order to make investment decisions for portfolio-building, the prediction threshold level of 0.8 would be about the right level. The 0.8 level is selected, because that gives an acceptable level of false positives, i.e. those low yields wrongly predicted as high yields.

The *Provinces A* geographical form of collapse gives a slightly better prediction than that for *Provinces B*. Therefore, the simplified model for *Provinces A* is preferred for the selection of premises in the building of a banking-hall investment property portfolio.

9.3 Analysis of Variance

9.3.1 Introduction

Sections 8.8.6 and 8.10 identified Analysis of Variance (ANOVA) as a worthwhile method of analysis to apply to the bank premises' sale-and-leaseback data. In ANOVA the dependent variable, *Yield*, is numeric rather than a factor as it was in the log-linear and logistic analyses. ANOVA is underpinned, however, by the assumption that the data have homogeneous variance and are normally distributed. Since the section 8.8.6 established that the available data do not have homogeneity of variance and show departures from normality, any models fitted using ANOVA will not be optimal in the least squares sense. In particular non-constant variance coupled with lack of balance in the data can lead to actual significance levels being quite different from the values assumed based on homoscedasticity (Wetherill, 1981: 264-266).

The lack of normality and constant variance means that when choosing which terms to eliminate from the model, it is possible to make some mistakes. However, the resultant model may still give some useful predictions. A model can be useful without being optimal.

Hence, a non-optimal ANOVA model can be tested to ascertain if it is more useful than the logistic model.

9.3.2 The Model for the *Provinces A* and the *Provinces* B datasets

The analysis of variance model for the current data set with the 4 factors *Bank* (B), *Lot Size* (L), *Province* (P) and *Time* (T) takes the form:

$$
\begin{aligned}
Y_i \quad = \quad & \lambda && \text{(1 constant term)} \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} && \text{(4 first order terms)} \\
& + \lambda_{BLjk} + \lambda_{BPjl} + \dots\dots\dots + \lambda_{PTlm} && \text{(6 second order terms)} \\
& + \lambda_{BLPjkl} + \lambda_{BLTjlm} + \dots\dots\dots + \lambda_{LPTklm} && \text{(4 third order terms)} \\
& + \lambda_{BLPTjklm} && \text{(1 fourth order term)}
\end{aligned}
$$

where Y_i is the predicted yield for the i'th case and

j, k, l and m represent the levels of the corresponding factors.

(9.10)

9.3.3 Analysis and Results

9.3.3.1 Analysis and Results of Provinces A

The saturated ANOVA model, comprising all the terms up to and including the fourth order, is shown in the equation 9.10. Backward elimination from the saturated model using the Akaike Information Criterion (AIC) suggested that no terms should be removed from the model. Therefore, the fourth order term was tested explicitly on an F-ratio test. F-ratios are used to test terms in ANOVAs (Field, 2005: 323-325). The F-ratio test showed that the fourth order term was not significant at the 10 per cent level. Therefore, the fourth order term was removed from the model before backward elimination was continued.

That produced a model with seven terms in the predictor variables:

$$
\begin{aligned}
Y_i \quad = \quad & \lambda \\
& +\lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{Tm} \\
& +\lambda_{BLjk} + \lambda_{BPjl} + \lambda_{BTjm}
\end{aligned}
\tag{9.11}
$$

These comprised the four first order terms, *Bank, Lot Size, Time* and *Provinces A*, and the following three second order interactive terms:

- Bank*Lot Size
- Bank*Time
- Bank*Provinces A

The significance of the three second order interactions were then tested explicitly. Two interactions, *Bank*Lot Size* and *Bank*Provinces A*, were significant at well beyond the 0.1 per cent level. *Bank*Time* was shown to be significant at the 5 per cent level and almost at the 1 per cent level. These three interactions were, therefore, retained in the model.

The output from the ANOVA with respect to the time model for Provinces A produced an Analysis of Deviance table which is reproduced in Table 9.21.

Table 9.21 Analysis of Deviance Table for the *Provinces A* ANOVA Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Bank	2	517.82	258.91	372.6232	< 2.2e-16	***
Lot Size	1	38.29	38.29	55.1028	3.467e-13	***
Time	1	346.74	346.74	499.0315	<2.2e-16	***
Provinces A	3	35.68	11.89	17.1189	1.007e-10	***
Bank:Lot Size	2	15.23	7.61	10.9580	2.074e-05	***
Bank:Time	2	6.41	3.21	4.6129	0.0102388	*
Bank: Provinces A	6	16.52	2.75	3.9623	0.0006637	***
Residuals	675	467.62	0.69			

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

The linear terms *Bank, Lot Size, Time* and *Provinces A* were all found to be significant at well beyond the 0.1% level in the F-test. The second order terms *Bank:Lot Size* and

Bank:Provinces A were found to be significant at well beyond the 0.1% level. The second order term *Bank:Time* was found to be significant at the 5% level.

The model itself is shown by the reduced Analysis of Variance table (Table 9.22) to be significant at well beyond the 0.1% level (F is 82.686 on 17 and 673 degrees of freedom).

Table 9:22 Significance of Fitted Model for *Provinces A* ANOVA Model

```
Model 1: yield ~ 1

Model  2:  yield  ~  bank  +  lotsize  +  time  +  provincesa  +
bank:lotsize + bank:time + bank:provincesa


      Resid.Df  Res.Sum of Sq  Df  Sum of Sq    F    Pr(>F)

Model 1   690      1444.32
Model 2   673       467.62     17   976.7    82.686 < 2.2e-16
```

The model was then checked for its goodness of fit. This was done by making a plot of the residuals against the fitted values to check for systematic changes in variance. The residuals are the difference between the actual yield and the fitted yield. Plotting the residuals against the fitted yield will mainly show if there is any trend in the variance of the residuals as the fitted value increases.

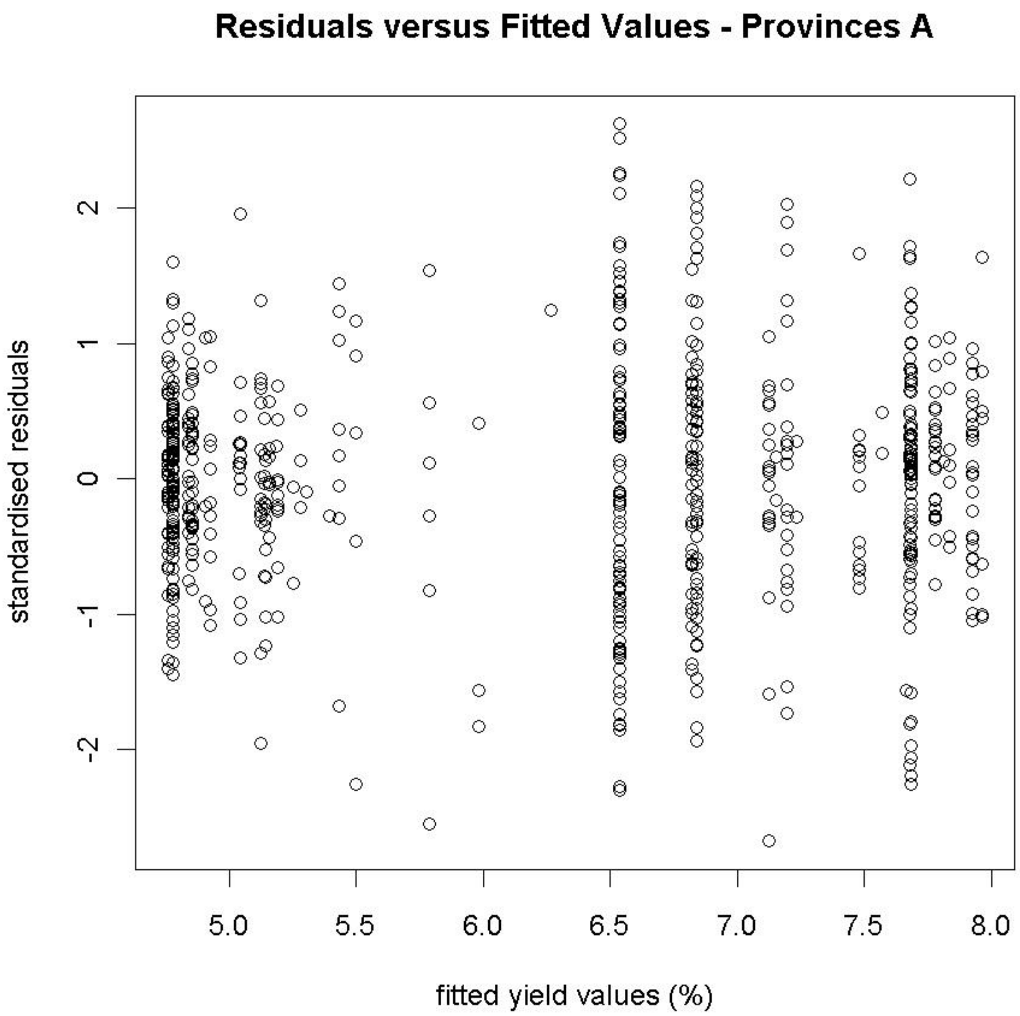


Figure 9.14 Residuals versus Fitted Values - *Provinces A*

A visual examination of the residual versus fitted values plot suggested that there were no clear systematic changes in variance. The residuals versus fitted values plot showed two distinct clusters that suggested that, in each of them, the spread of the residuals (standard deviation) decreased as the fitted yield values increased. Respectively, the two clusters were grouped above and below the fitted value of about 6.35 per cent, which corresponds to the low point in the *Yield* distribution curve in Figure 8.6. Clearly, the commonly used variance stabilizing transformations (e.g. square root and log) are unlikely to be effective in this case.

A Q-Q plot provides an illustration of deviation from normality. Q-Q plots plot the quantiles of the fitted values against the theoretical quantiles of the test distribution. The expected positions are represented by a straight line whilst the observed values are represented by individual dots made on the plot (Field). Where the data are normally distributed, these dots will fall exactly on the diagonal straight line, since the values coincide with those to be expected from a normally distributed dataset (Field). Hence, any divergence between the dots and the line is indicative of deviation from normality.

A Q-Q plot of the residuals (Figure 9.15) was made to give a visual check for normality. The tail ends are off the fitted line, which suggests that the distribution has a greater peak than a normal distribution would have. Hence, the distribution is leptokurtic. ANOVA tests are known to be particularly sensitive to this type of departure from normality (Scheffé, 1959; Wetherill, 1981: 18-20).

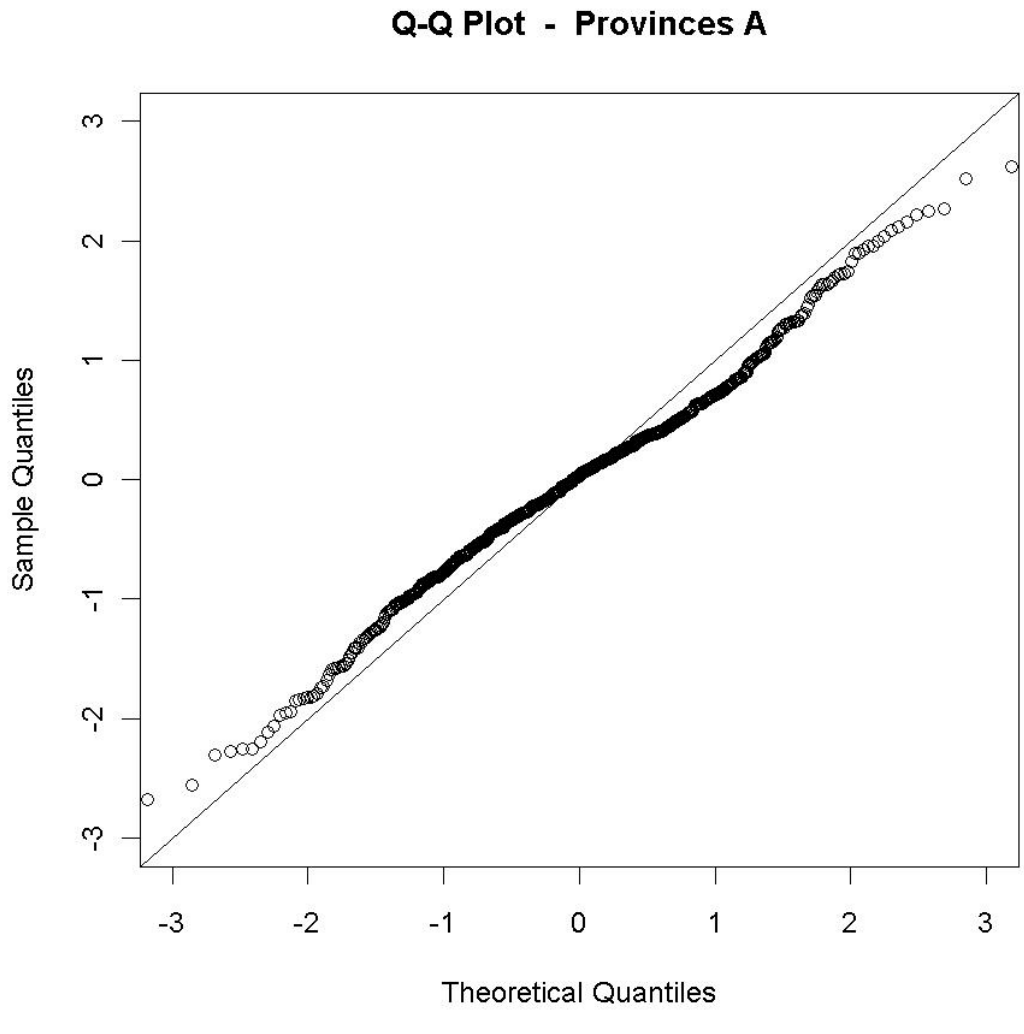


Figure 9.15 Q-Q Plot – *Provinces A*

A confusion matrix was derived from the plot (similar to that in Figure 9.16) of Predicted Yield against Actual Yield. A threshold of 6.35 per cent was applied to both yield values. This divided the plot into four quadrants from which a confusion matrix could be generated. The confusion matrix is reproduced in Table 9.23.

Table 9.23 Confusion Matrix of Provinces A ANOVA Model before cross-validation: Threshold value of 6.35%

		Predicted	
		Low Yield	High Yield
Actual	Low Yield	287	86
	High Yield	13	305

Cross-validation was performed using random splits of ten without replication. An almost identical confusion matrix was produced after cross-validation and is reproduced in Table 9.24. As in the logistic regression, the cross-validation results for the ANOVA are in general no worse than before cross-validation.

Table 9.24 Confusion Matrix of Provinces A ANOVA Model after Cross-validation: Threshold value of 6.35%

		Predicted	
		Low Yield	High Yield
Actual	Low Yield	287	86
	High Yield	12	306

In order to look at the practical usefulness of the results, the predicted yield was plotted against the actual yield. This plot is shown in Figure 9.16. The horizontal and vertical lines in the figure represent a threshold of 6.35 per cent. The confusion matrix is derived from the number of points in each quadrant.

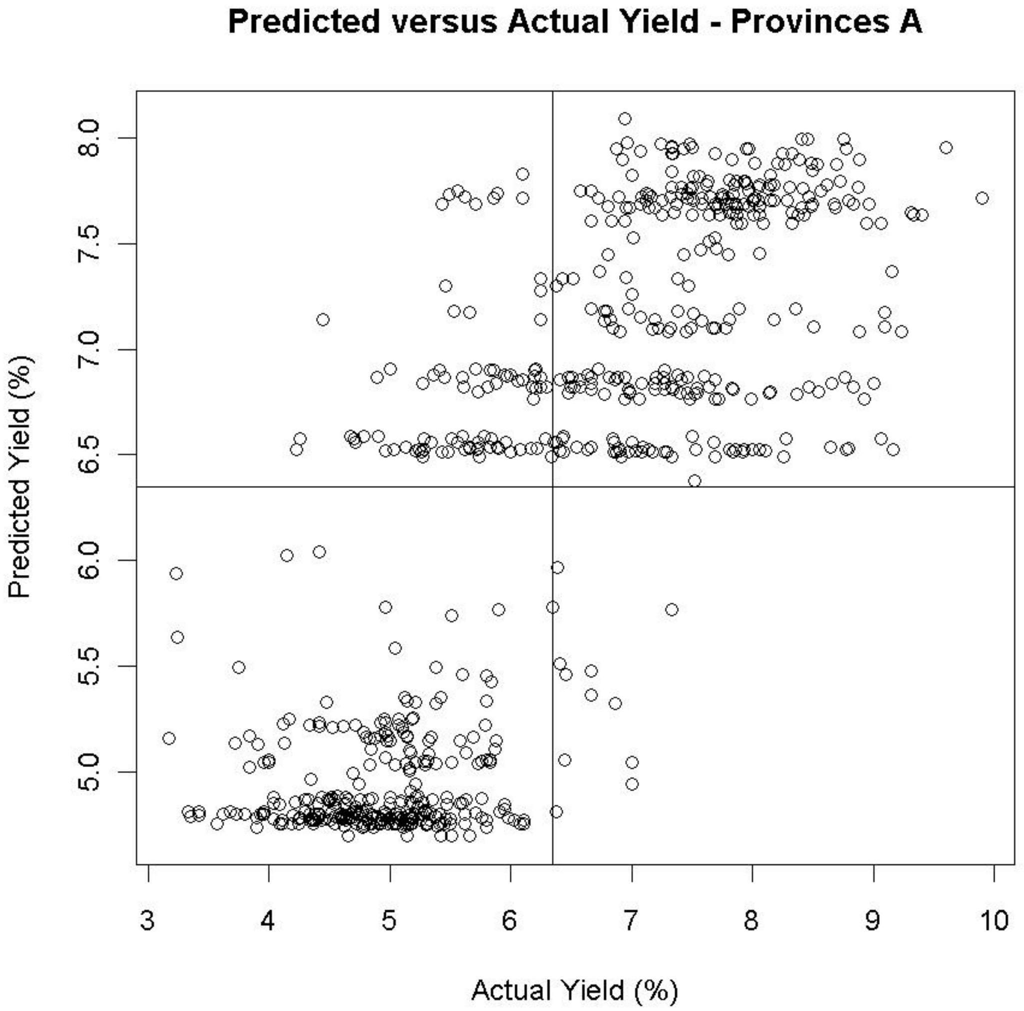
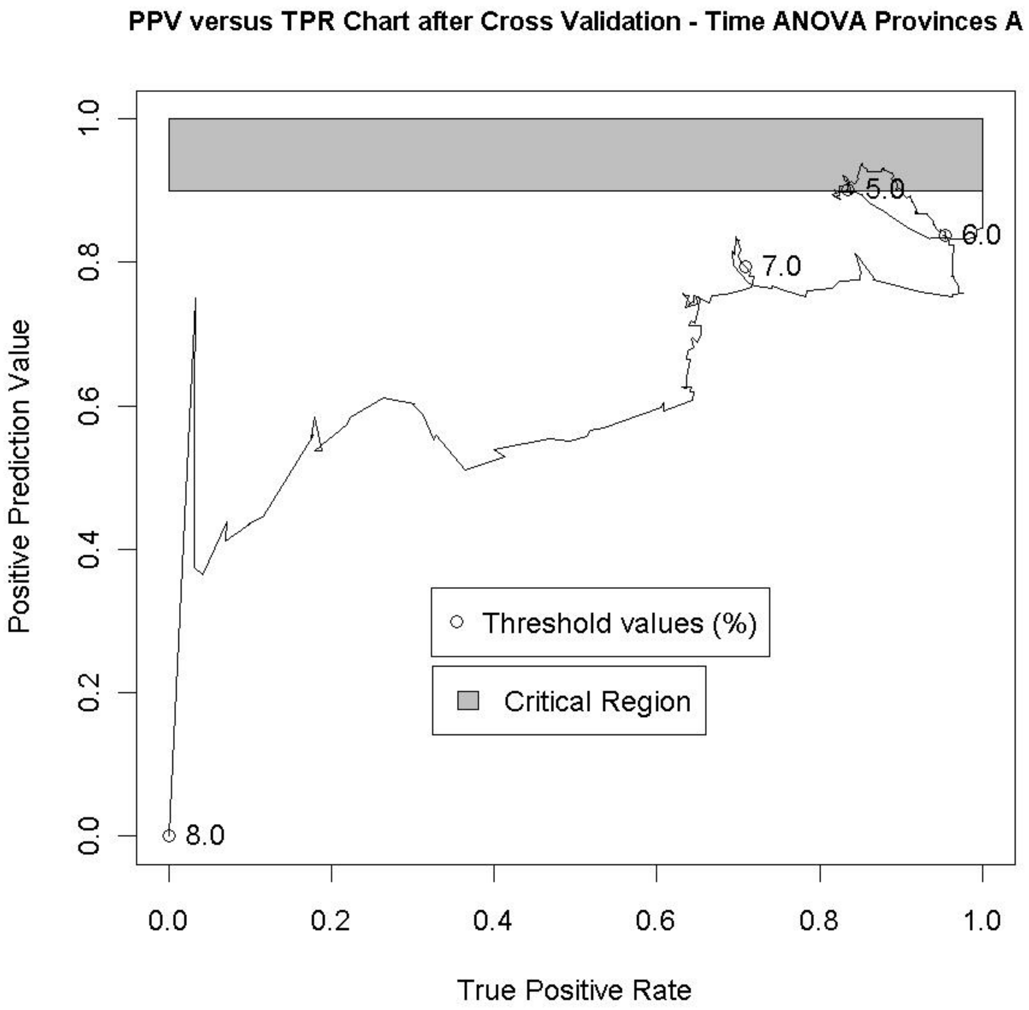


Figure 9.16 Predicted versus Actual Yield – *Provinces A*

It is exceedingly difficult to identify a good threshold from this graph.

Figure 9.17 shows the Predicted Positive Value versus True Positive Rate after cross-validation with respect to the time ANOVA model for *Provinces A*.



9.17 **PPV versus TPR Chart after Cross-Validation – *Time* ANOVA *Provinces A***

The Predictive Positive Value versus True Positive Rate chart in Figure 9.17 is very erratic with the line of the plots looping in several places. Moreover, apart from a very small area within one of the loops, none of the selected predictive positive rate of ninety per cent or more for high yield was achieved.

Figure 9.17 can be compared and contrasted with Figure 9.5 (logistic regression) which also shows a Predicted Positive Value versus True Positive Rate and which chart confirms that 0.8 appears to be a good threshold.

Figure 9.17 demonstrates that the ANOVA model for *Provinces A* does not provide a meaningful basis for selecting the threshold required in order to make useful predictions for investing in retail bank premises.

9.3.3.2 Analysis and results of Provinces B

The analysis of the *Provinces B* dataset using ANOVA was undertaken in the same sequence as that for *Provinces A*, using the Akaike Information Criterion (AIC) based upon backward elimination. Then the same plots were made as in the other dataset in order to try to ascertain the usefulness of the model. Again, backward elimination from the saturated model was undertaken.

Backward elimination from the saturated model using the Akaike Information Criterion (AIC) provided a model with five terms in the predictor variables. These are expressed in the following equation:

$$
\begin{aligned}
Y_i \quad = \quad & \lambda \\
& + \lambda_{Bj} + \lambda_{Pl} + \lambda_{Tm} \\
& + \lambda_{BPjl} + \lambda_{BTjm}
\end{aligned}
\qquad (9.12)
$$

The model omitted *Lot Size* altogether. Therefore, the linear terms comprised *Bank, Time* and *Provinces B*. It also produced two second order terms, which were *Bank*Time* and *Bank*Provinces B*. The two second order terms were tested explicitly on F-ratio tests. These tests showed that the *Bank*Time* interaction was significant at beyond the 0.1 per cent level and that the *Bank*Provinces B* term was significant at the 1.0 per cent level. Therefore, both terms were retained in the model.

Table 9.25 Analysis of Deviance Table for the *Provinces B* ANOVA Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Bank	2	517.82	258.91	355.1011	< 2.2e-16	***
Time	1	384.80	384.80	527.7589	<2.2e-16	***
Provinces B	3	16.34	5.45	7.4705	6.326e-05	***
Bank:Time	2	17.07	8.53	11.7053	1.006e-05	***
Bank: Provinces B	6	15.40	2.57	3.5208	0.001942	**
Residuals	676	492.89	0.73			

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

The linear terms *Bank, Time* and *Provinces B* were all found to be significant at well beyond the 0.1% level in the F-test. The second order term *Bank:Time* was found to be significant at well beyond the 0.1% level. The second order term *Bank:Provinces B* was found to be significant at the 1% level.

The model itself is shown by the reduced Analysis of Variance table (Table 9.26) to be significant at well beyond the 0.1% level (F is 93.208 on 14 and 676 degrees of freedom).

Table 9:26 Significance of Fitted Model for *Provinces B* ANOVA Model

```
Model 1: yield ~ 1

Model 2: yield ~ bank + time + provincesb + bank:time +
bank:provincesb


      Resid.Df  Res.Sum of Sq   Df  Sum of Sq   F     Pr(>F)

Model 1   690      1444.32
Model 2   676        492.89     14   951.43   93.208 < 2.2e-16
```

The *Provinces B* ANOVA model has a different structure from the *Provinces A* one. This is unlike the case in the logistic regression where *Provinces A* and *Provinces B* produced the same simplified model. This shows that there is less stability in the ANOVA than in the logistic regression in terms of model structure.

A confusion matrix was again derived from the plot of Predicted Yield versus Actual Yield using a threshold of 6.35 per cent. This is shown in Table 9.27.

Table 9.27 Confusion Matrix of *Provinces B* ANOVA Model produced before and after Cross-validation: Threshold value of 6.35%

		Predicted	
		Low Yield	High Yield
Actual	Low Yield	287	86
	High Yield	13	305

A visual examination of the plot of the Residuals versus Fitted Values for *Provinces B*, shown in Figure IX.1 in Appendix IX, suggested that there were two distinct clusters. Again, the chart suggested that the commonly used variance stabilizing transformations would be unlikely to be effective. Again, the Q-Q plot for *Provinces B*, shown in Figure IX.2 in the Appendix IX, indicated that the distribution is leptokurtic.

Cross-validation was performed using random splits of ten without replication. The confusion matrices produced before and after cross-validation were identical. In order to look at the practical usefulness of the results, the predicted yield was plotted against the actual yield. This plot is shown in Figure IX.3 in Appendix IX. The horizontal and vertical lines in the figure represent a threshold of 6.35 per cent. The confusion matrix is derived from the number of points in each quadrant.

The Predictive Positive Value versus True Positive Rate chart in Figure IX.4 in Appendix IX is very erratic with the line of the plots looping in several places. Moreover, apart from a very small area within one of the loops, none of the selected predictive positive rates of ninety per cent or more for high yield was achieved.

Figure IX.4 in Appendix IX can be compared and contrasted with Figure 9.11 (logistic regression) which also shows a Predicted Positive Value versus True Positive Rate and which chart confirms that 0.8 appears to be a good threshold.

Figure IX.4 demonstrates that the ANOVA model for *Provinces B* does not provide a meaningful basis for selecting the threshold required in order to make useful predictions for investing in retail bank premises.

The output of the charts derived from the confusion matrix suggests that the selection of a (prediction) threshold is as intractable for the *Provinces B* dataset as it was for the *Provinces A* dataset. The model produced for the Provinces B data is very similar, if not a little worse, in terms of interpreting the data. There was no clear, obvious choice of a good threshold.

9.3.4 Conclusions to the ANOVAs

ANOVA models were fitted to both the *Provinces A* and *Provinces B* data. For Provinces A the model comprised seven terms in the predictor variables. The four first order effects plus the three second order interactions including Bank. For *Provinces B*, the model comprised five terms in the predictor variables. Lot size was omitted both as a linear term and as an interaction. In *Provinces B*, the *Bank*Time* and the *Bank*Provinces B* interactions were found to be significant.

In both cases, the distribution of predicted versus actual yield was such that the problem of selecting a suitable threshold proved to be intractable. The results were erratic and no feasible criteria for a good threshold could be found.

9.4 Conclusions to Advanced Analyses using calendar data

9.4.1 Logistic Regression

- For both Provinces A and Provinces B the logistic regression analysis produced a simplified model comprising five terms in the predictor variables. These were the four linear terms plus the Bank*Time interaction.

- In practical terms the simplified models proved to be as useful as the more complex basic models from which they were derived.

- In both cases, a probability threshold at or above 0.8 was found to give sufficiently accurate predictions for application for investment decisions in portfolio-building.

9.4.2 ANOVA

- For *Provinces A*, the ANOVA produced seven terms in the predictor terms. These were the four first order terms plus the three second order interactions involving Bank.

- For *Provinces B*, the ANOVA produced a model comprising five terms in the predictor variables. These were Bank, Time and Provinces as first order terms plus the second order interactions of *Bank*Time* and *Bank*Provinces B*.

- In both cases, the distribution of predicted versus actual yield was such that the problem of selecting a suitable threshold proved to be intractable. The results were erratic and no feasible criteria for a good threshold could be found.

9.5 Final observation

Binary logistic regression proved to produce a more useful model than ANOVA. The simplified model produced the same results as the basic model generated by binary logistic regression.

The logistic regression models for both datasets included the four main independent variables identified by the qualitative study. This is qualified inasmuch as *Province* (or super-region), rather than GOR region, is shown to be a factor. These results, in turn, had been supported by the literature review. The *Provinces A* ANOVA model, although shown not to provide a meaningful basis for selecting the threshold required in order to make useful predictions, also suggested that these four independent variables influenced *Yield*.

10 Advanced Analyses II: Original dataset incorporating index data

10.1 Introduction

The previous analyses in Chapters 8 and 9 have used *Time* as one of the independent variables. The reason for that is that time is representative of fluctuations in the macro economic cycle. Property performance is influenced by those fluctuations. This chapter examines the effect of substituting an appropriate index for *Time*. The reason for so doing is to establish whether the substitution of an appropriate index improves the predictive accuracy of models included in the predictive framework adopted by this study.

There are indices that relate specifically to the performance of real estate. Such indices relate to different geographical locations and may relate to different sectors of property. They also measure property performance in different ways. Therefore, it is imperative to select the index that is most relevant to this study.

The Investment Property Databank (IPD) publishes a number of property indices for different sectors of property across different geographical locations around the world. It does publish a monthly index of initial yield for retail property across the United Kingdom (Investment Property Databank, (2008). The data provided by this index covers all of the study period. The IPD retail property index for the United Kingdom does not provide data specifically related to banking-halls. However, since banking-halls have been shown to be a sub-type of retail property in the Theoretical Perspectives Chapter (Chapter 4), the IPD United Kingdom retail property index appears to be the most appropriate one to substitute for *Time*.

Property investment yields may be affected by both market risk and specific risk (Isaac and Steley, 2000; Fraser, 2004). Market risk arises out of fluctuations in the macro-economic cycle over time (Fraser, 2004: 158). Isaac and Steley (2000: 47-49) argue that in deciding on the anticipated returns that they will accept for any given investment, investors will consider appropriate indices. Hence, they may consider the rates of index-linked gilt edged securities or they may look to one of the Government's inflation indices (Isaac and Steley). A study into any given type of investment should use an index most appropriate to that type of investment. The Investment Property Databank (IPD) United Kingdom Retail Property Index covering the period of the current study has been made available to this research. Therefore, this latter index has been adopted as the most appropriate one to use in this study.

As when *Time* was a predictor variable, when it is substituted by the *Index* variable, there are two bases of models in the framework. These two further bases of models are as follows:

3. The main intention is to predict *Yield Group*, which is what the majority of investors want to know. Logistic regression is a statistical procedure that would enable this.

4. If an ANOVA model can be developed, it would allow investors to go a step further in being able to assess the yields and then the value at which they should bid by capitalizing rents at the appropriate yield.

The original data made available by the auction houses about British banking-halls, sold at auction as investments, provided a data variable. The IPD initial yield index for United Kingdom retail property is produced on a monthly basis. The data variable was used to identify the appropriate index value for each case in the dataset. This chapter therefore examines logistic regression and ANOVA models using the IPD initial yield retail property *Index* as an independent variable.

10.2 Logistic regression

10.2.1 Introduction to logistic regression

The same procedures were adopted for the logistic regression in this section as those that were used when *Time*, rather than *Index*, had been an independent variable.

Whereas *Time* was a factor, or categorical variable, *Index* is a continuous variable. Therefore, the *Index* variable does not show factor levels in the equations following.

10.2.2 The Model for the *Provinces A* and *Provinces B* datasets

The full, or saturated, binary logistic model for the two datasets with dependent variable *Yield Group* (Y) has as predictor variables the three factors, *Bank*, *Lot Size* and *Provinces*, and the numeric variable, *Index*, and takes the form:

$$
\begin{aligned}
\log\left(\frac{p_i}{1-p_i}\right) = \;& \lambda && \text{(1 constant term)} \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_I I_i && \text{(4 first order terms)} \\
& + \lambda_{BLjk} + \lambda_{BPjl} + \dots\dots\dots + \lambda_{PIl} I_i && \text{(6 second order terms)} \\
& + \lambda_{BLPjkl} + \lambda_{BLIjl} I_i + \dots\dots\dots + \lambda_{LPIkl} I_i && \text{(4 third order terms)} \\
& + \lambda_{BLPIjkl} I_i && \text{(1 fourth order term)}
\end{aligned}
\tag{10.1}
$$

Where:

- p_i is the predicted probability that the i'th case belongs to the high yield group,
- I_i is the index value for the i'th case and
- *j, k* and *l* represent the categories of the factors *B, L,* and *P* respectively.

Equation 10.1 can be re-arranged to give:

$$p_i = \exp(\, f(\lambda)\,) \,/\, (1 + \exp(\, f(\lambda)\,)\,)$$

where

$$
\begin{aligned}
f(\lambda) = \;& \lambda \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_I I_i \\
& + \lambda_{BLjk} + \lambda_{BPjl} + \dots\dots\dots + \lambda_{PIl} I_i \\
& + \lambda_{BLPjkl} + \lambda_{BLIjl} I_i + \dots\dots\dots + \lambda_{LPIkl} I_i \\
& + \lambda_{BLPIjkl} I_i
\end{aligned}
\tag{10.2}
$$

10.2.3 Analysis

10.2.3.1 Analysis of Provinces A

Backward elimination from the saturated model gave a basic model incorporating the IPD index (equation 10.3) comprising five terms in the predictor variables:

$$
\begin{aligned}
f(\lambda) \quad = \quad & \lambda \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_I I_i \\
& + \lambda_{BIj} I_i
\end{aligned}
\qquad (10.3)
$$

The significance of the second order term Bank*Index was explicitly tested. The results of that test are reproduced in Table 10.1 which shows the term not to be significant at the ten per cent level. Therefore, it was removed from the model before the continuation of the backward elimination process.

Table 10.1 Explicit Test of Bank:Index Term

```
Model 1: yield ~ bank + index + lotsize + provincesa + bank:index

Model 2: yield ~ bank + index + lotsize + provincesa

     Resid.Df  Resid.Dev  Df   Deviance   P(>|Chi|)
1       681     365.58
2       683     370.03    -2    -4.4486     0.1081
```

Backward elimination following the removal of the Bank*Index term produced a model containing three terms in the predictor variables.

The resultant simplified index model (known as the *Index* model) produced the following equation:

$$
\begin{aligned}
f(\lambda) \quad = \quad & \lambda \\
& + \lambda_{Lk} + \lambda_{Pl} + \lambda_I I_i
\end{aligned}
\qquad (10.4)
$$

These were the three first order terms: Index, Lot Size and Provinces A. All three of these first order terms were shown to be significant at well beyond the 0.1 per cent level. The significance of these three terms is shown in the Analysis of Deviance table reproduced in Table 10.2.

Table 10.2 Analysis of Deviance Table for the Provinces A Model including Index

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	
NULL			690	953.86		
Index	1	516.03	689	437.83	< 2.2e-16	***
Lot size	1	10.49	688	427.34	0.001199	**
Provinces A	3	53.99	685	373.35	1.127e-11	***

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

Index is shown by Table 10.2 to have much greater explanatory power than Lot Size and Provinces. The table of coefficients (Table 10.3) indicates that the higher the index, the more likely a banking-hall is to be in the high yield group. The large Lot Size has a significantly different effect from the small Lot Size. Larger lots are more likely to result in higher yields. The Midlands and South have a significantly different effect from London as far as Provinces is concerned. The yields are shown to be higher in the north of England and the Celt provinces.

In an apparent contrast to the findings of the qualitative study, the term *Bank* has been excluded from the model.

The *Bank* and the *Bank:Time/Index* terms were present in both the simplified and the index models according to the Akaike Information Criterion (AIC). When the terms were tested explicitly, the levels of statistical significance were quite different in the simplified and index models. Due to these differences, the practical significance of the terms was explicitly tested in both cases and the terms were found to be practically significant in the simplified model (comprising the *Time* variable), but not the index model. In the simplified model, the bank term and the *Bank:Time* interaction were both practically and statistically significant. In the indexed model, they were the most statistically significant terms to be omitted from the practical model, because they lacked sufficient practical significance. The main reason why *Bank* falls out when the *Bank:Index* interaction is omitted would appear to be that *Index* has much greater explanatory power than *Bank*.

Table 10.3 Coefficients for Index Model for Provinces A

	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	-23.4374	2.4067	-9.738	< 2.2e-16	***
Index	3.6291	0.3499	10.372	< 2.2e-16	***
Lot Size - Large	1.1638	0.3307	3.519	0.000433	***
Provinces A - Midlands	-1.3574	0.4999	-2.715	0.006624	**
Provinces A - South	-2.1574	0.4724	-4.567	4.95e-06	***
Provinces A - Celt	0.9067	0.7566	1.199	0.230722	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The Coeffficients for Index Model for Provinces A show that *Index* is highly significant. The large *Lot Size* is significantly different from the small *Lot Size*. The Midlands and the south of England are significant in relation to the north of England province. However, the Celt province is not significant in relation to the north of England. These results support the hypotheses that the dependent variable *Yield* is affected by the independent variables *Lot Size*, *Region* (but as a collapsed into *Provinces A*) and *Index*.

Confusion matrices were produced for the model both before and after cross-validation. The confusion matrix produced by the model before cross-validation is reproduced in Table 10.4.

Table 10.4 Confusion Matrix of Provinces A Index Model: p = 0.5

		Actual Yield	
		Low Yield	High Yield
Predicted	Low Yield	307	39
	High Yield	65	280

After cross-validation, the index model for Provinces A produced a probability bar chart which is reproduced in Figure 10.1.

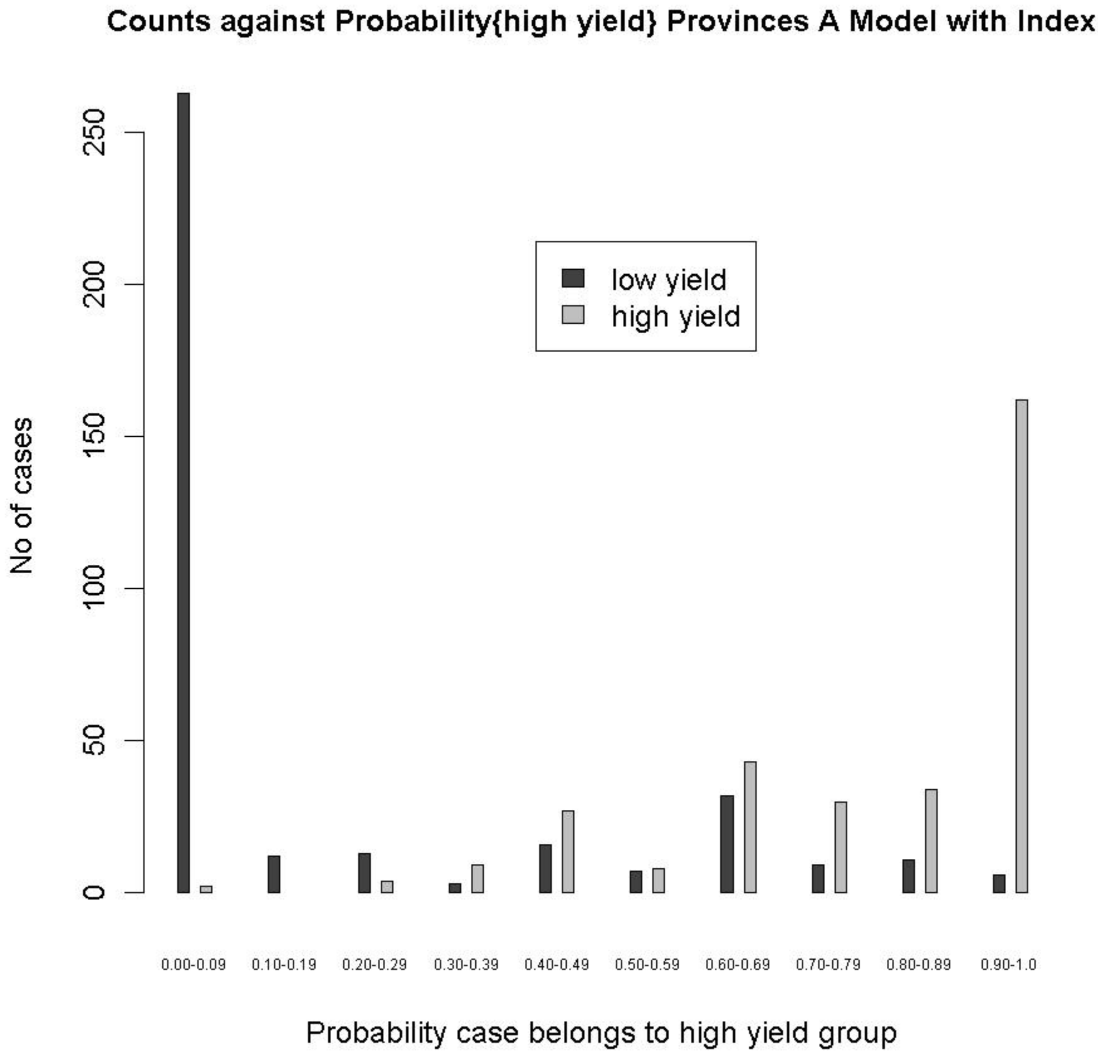


Figure 10.1 Probability bar chart for the index model for *Provinces A* after cross-validation

This gives rise to the following confusion matrix in Table 10.5 which is very similar to the one in Table 10.4 for the model before cross-validation.

Table 10.5 Confusion Matrix of Provinces A Index Model after Cross-validation: p = 0.5

		Actual Yield	
		Low Yield	High Yield
Predicted	Low Yield	307	42
	High Yield	65	277

As for the *Time*-based logistic model, a ROC chart was produced for the index model for *Provinces A* following cross-validation. This can be used to identify the best threshold for predicting high yielding banking-halls. This ROC chart is shown in Figure 10.2.

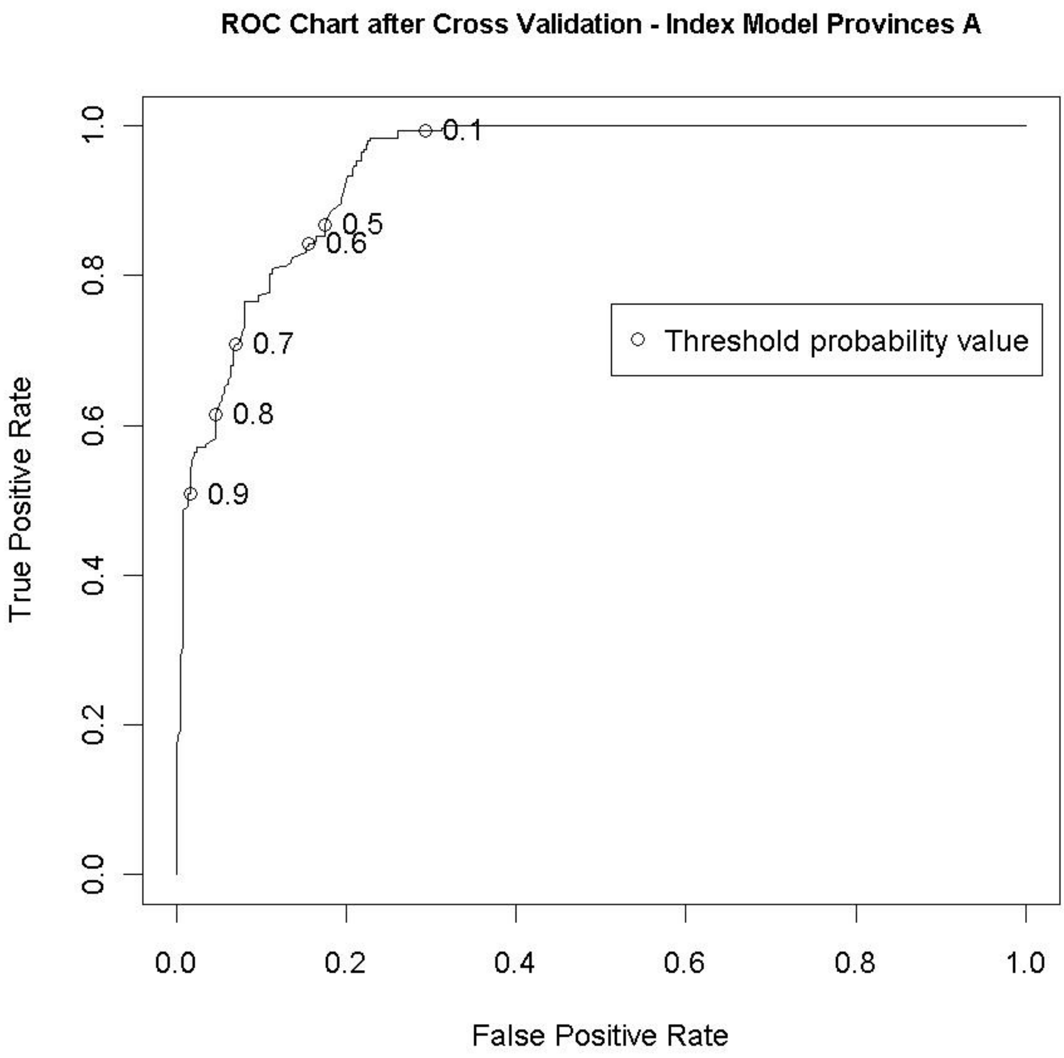


Figure 10.2 ROC Chart after Cross Validation – Index Model *Provinces A*

Examination of the predictions bar chart and the ROC chart suggest that for an acceptably low (0.1 or less) false positive rate, a threshold of 0.8 or more is required.

A plot of the predicted positive value versus the true positive rate was also made after cross-validation. This chart is shown in figure 10.3.

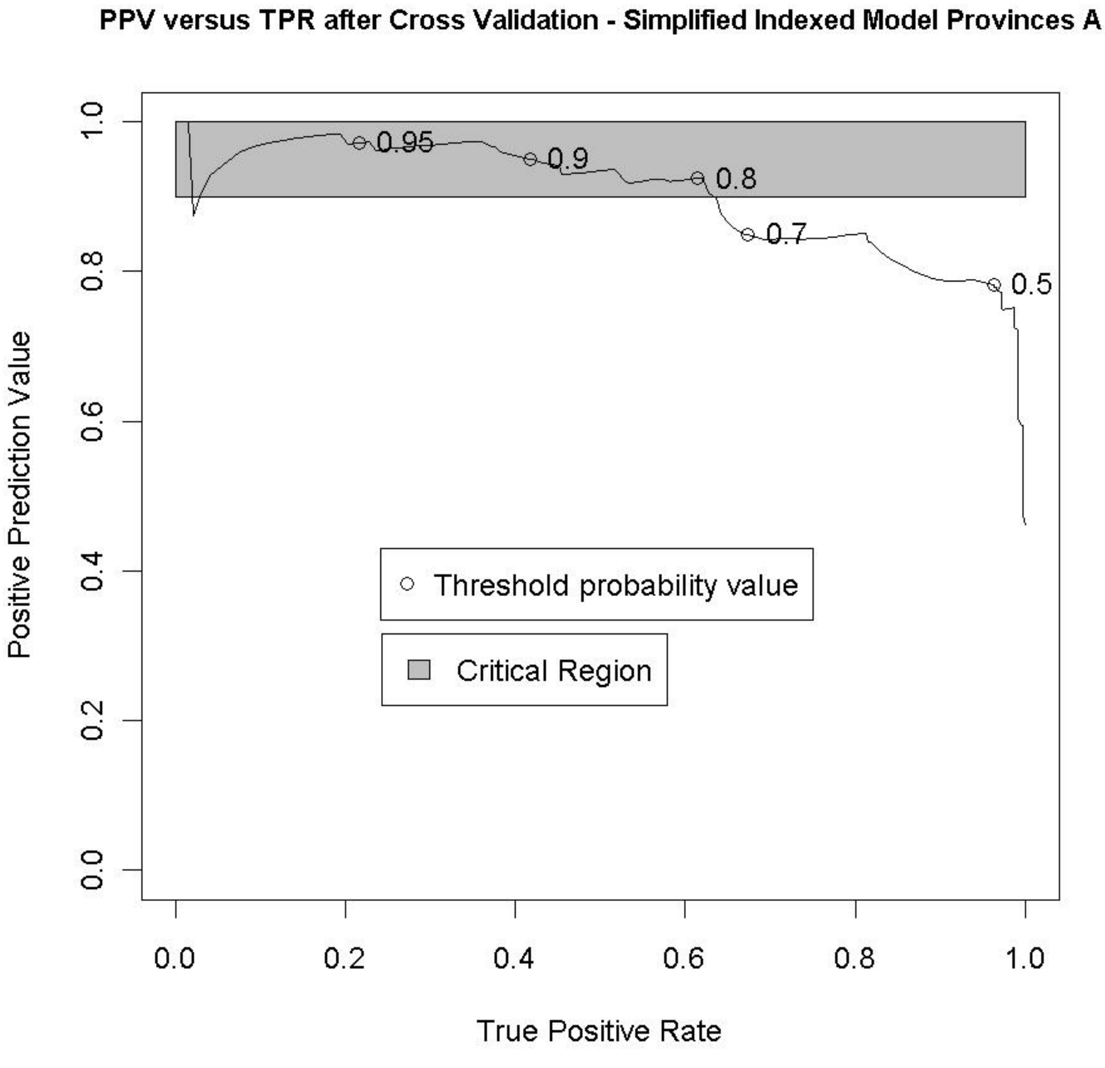


Figure 10.3 PPV versus TPR after Cross-Validation – Simplified Indexed Model *Provinces A*

The erratic behaviour of the graph in Figure 10.3 when the threshold is at an extremely high level very close to the maximum actual or predicted probability results from the very few actual or predicted cases which are above the threshold. The chart confirms, again, that 0.8 appears to be a good threshold.

Following the generation of the two ROC charts with respect to the basic and simplified indexed models for *Provinces A*, a further ROC chart was produced, comparing the predictive accuracy of the two models. This third ROC chart is shown in Figure 10.4. Clearly, this ROC chart suggests that there is very little difference between the predictive powers of the basic and simplified indexed models with respect to *Provinces A*. Indeed, an eyeball examination of the chart suggests that the simplified model may actually have marginally greater predictive power than the basic model. The simplified model is, therefore, preferred.

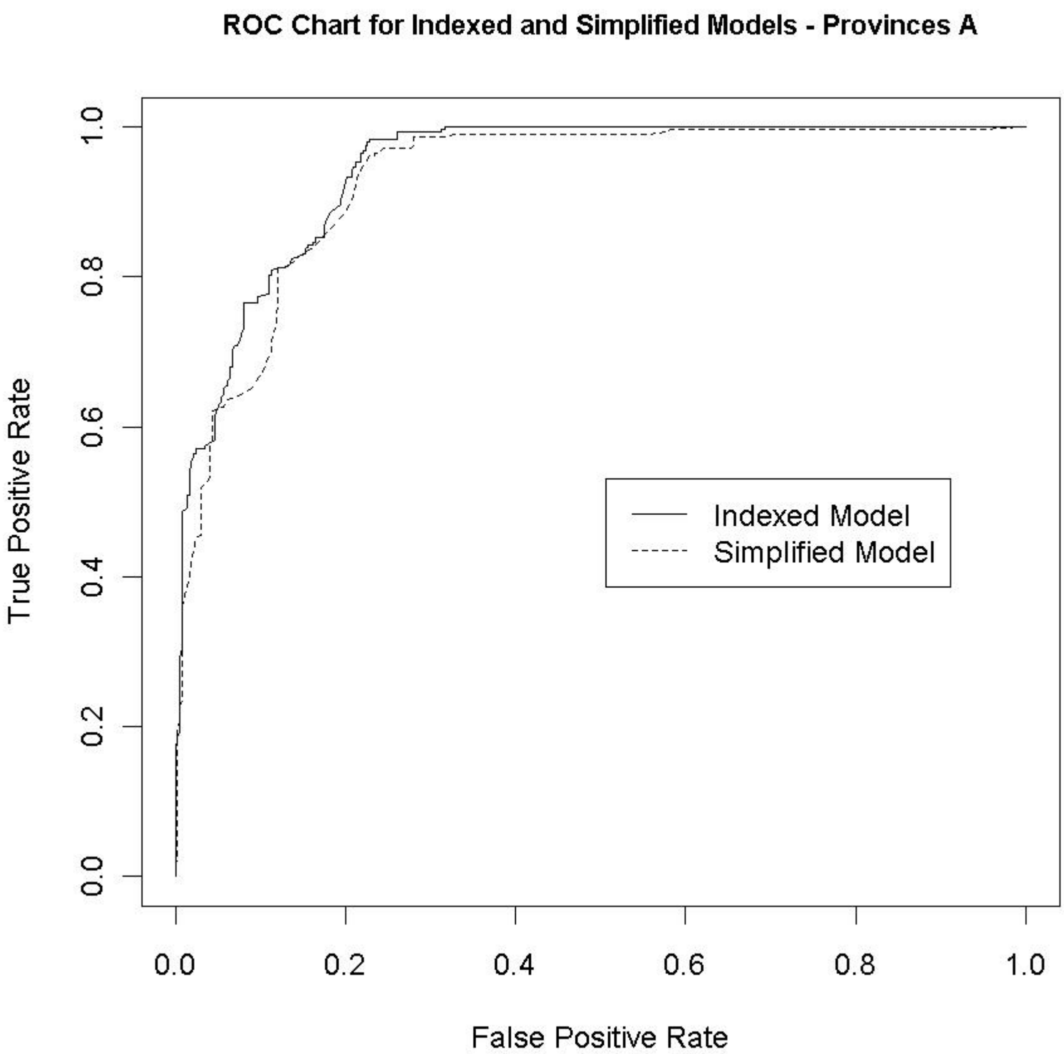


Figure 10.4 ROC Chart for Indexed and Simplified Models – *Provinces A*

10.2.3.2 Analysis of Provinces B

Backward elimination from the saturated model gave a basic model incorporating the IPD index (equation 10.5) comprising five terms in the predictor variables:

$$
\begin{aligned}
f(\lambda) \quad = \quad & \lambda \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{I} I_i \\
& + \lambda_{BIj} I_i
\end{aligned}
\qquad (10.5)
$$

The significance of the second order term *Bank*Index* was explicitly tested. The results of that test are reproduced in Table 10.6 which shows the term not to be significant at the ten per cent level. Therefore, it was removed from the model before the continuation of the backward elimination.

Table 10.6 Explicit Test of Bank:Index Term

```
Model 1: yield ~ bank + index + lotsize + provincesb + bank:index
Model 2: yield ~ bank + index + lotsize + provincesb
     Resid.Df   Resid.Dev    Df   Deviance   P(>|Chi|)
1       681      386.17
2       683      390.75      -2   -4.5814     0.1012
```

Backward elimination following the removal of the *Bank*Index* term produced a model containing three terms in the predictor variables.

The resultant simplified index model (known as the *Index* model) produced the following equation:

$$
\begin{aligned}
f(\lambda) \quad = \quad & \lambda \\
& + \lambda_{Lk} + \lambda_{Pl} + \lambda_{I} I_i
\end{aligned}
\qquad (10.6)
$$

These were the three first order terms: Index, Lot Size and Provinces B. All three of these first order terms were shown to be significant at well beyond the 0.1 per cent level. The significance of these three terms is shown in the Analysis of Deviance table reproduced in Table 10.7.

Table 10.7 Analysis of Deviance Table for the Provinces B Model including Index

	Df	Deviance	Resid. Df	Resid. Dev	P(>|Chi|)	
NULL			690	953.86		
Index	1	516.03	689	437.83	< 2.2e-16	***
Lot size	1	10.49	688	427.34	0.001199	**
Provinces B	3	33.04	685	394.30	3.163e-07	***

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

Index is shown by Table 10.7 to have much greater explanatory power than Lot Size and Provinces. The table of coefficients (Table 10.8) indicates that the higher the index, the more likely a banking-hall is to be in the high yield group. The large Lot Size has a significantly different effect from the small Lot Size. Larger lots are more likely to result in higher yields. Yields with respect to North Britain are significantly greater than those for other three *Provinces*.

Again, in an apparent contrast to the findings of the qualitative study, the term *Bank* has been excluded from the model.

The *Bank* and the *Bank:Time/Index* terms were present in both the simplified and the index models according to the Akaike Information Criterion (AIC). When the terms were tested explicitly, the levels of statistical significance were quite different in the simplified and index models. Due to these differences, the practical significance of the terms was explicitly tested in both cases and the terms were found to be practically significant in the simplified model (comprising the *Time* variable), but not the index model. In the simplified model, the bank term and the *Bank:Time* interaction were both practically and statistically significant. In the indexed model, they were the most statistically significant terms to be omitted from the practical model, because they lacked sufficient practical significance. The main reason why

Bank falls out when the *Bank:Index* interaction is omitted would appear to be that *Index* has much greater explanatory power than *Bank*.

Table 10.8 Coefficients for Index Model for Provinces B

	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	-25.19282	2.49175	-10.110	< 2.2e-16	***
Index	3.61880	0.34970	10.348	< 2.2e-16	***
Lot Size - Large	0.94433	0.32474	2.908	0.00364	**
Provinces B – Wales & SW	0.02356	0.31917	0.074	0.94115	
Provinces B - Midlands	0.52733	0.33597	1.570	0.11652	
Provinces B – North Britain	2.28065	0.48303	4.722	2.34e-06	***

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The Coeffficients for Index Model for Provinces B show that *Index* is highly significant. The large *Lot Size* is significantly different from the small *Lot Size*. As far as Provinces is concerned, the north of Britain is significantly different from the rest of the country. Table 10.8 clearly shows that the north of Britain is significantly different from London and the south-east of England. Comparisons between north Britain, the Midlands, and Wales and the south-west, based on Tables 10.8, indicate that the north of Britain is significantly different from the other two provinces at beyond the 10 per cent level. However, there was no significant difference between the other provinces. These results support the hypotheses that the dependent variable *Yield* is affected by the independent variables *Lot Size, Region* (as collapsed as *Provinces B*) and *Index*.

Confusion matrices were produced for the model both before and after cross-validation. The confusion matrix produced by the model before cross-validation is reproduced in Table 10.9. At the p = 0.5 threshold, it showed a fairly decent model, because the true positive rate was 79.6 per cent.

Table 10.9 Confusion Matrix of Provinces B Index Model: p = 0.5

		Actual Yield	
		Low Yield	High Yield
Predicted	Low Yield	295	18
	High Yield	77	301

After cross-validation, the index model for *Provinces B* produces a probability bar chart which is reproduced in Figure 10.5.

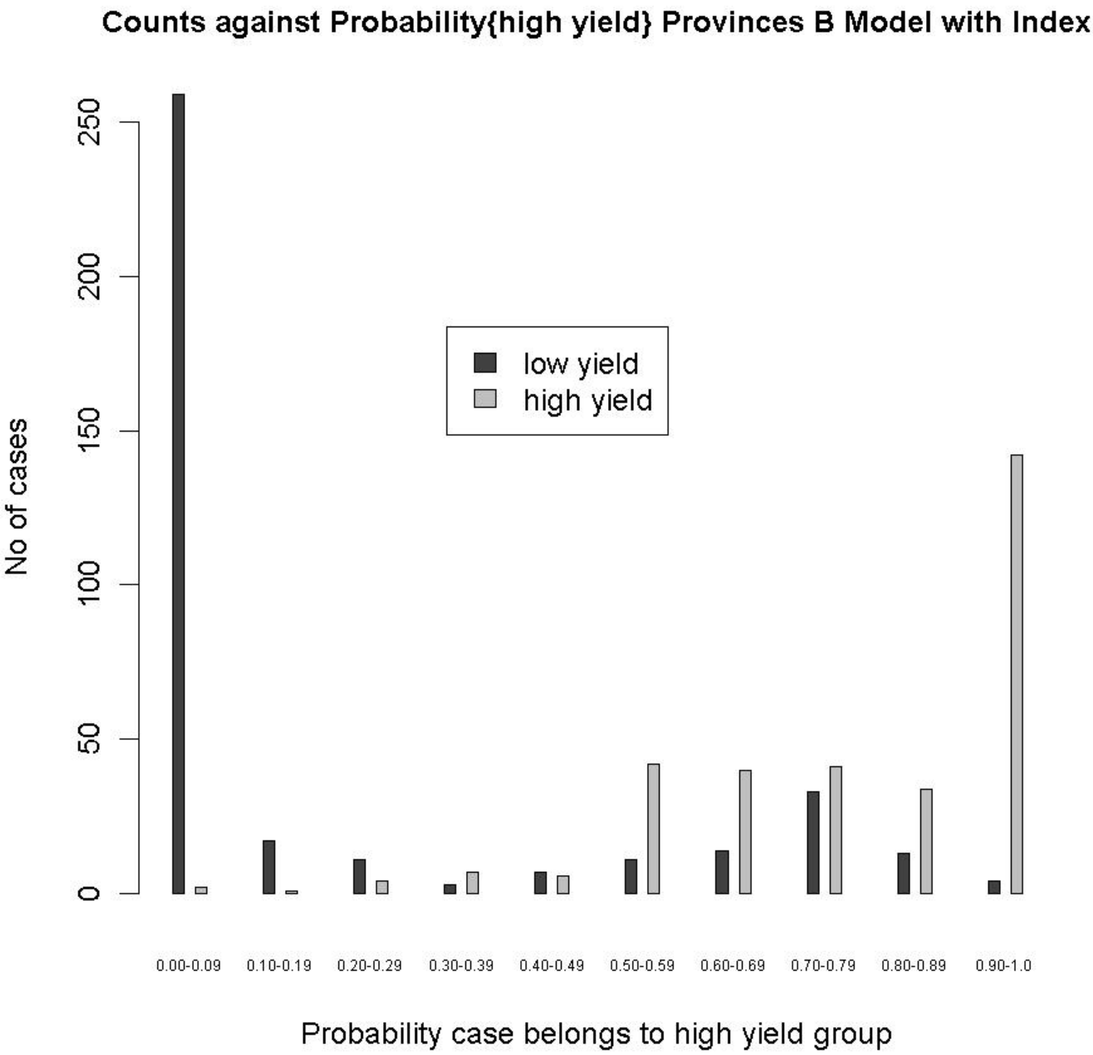


Figure 10.5 Probability bar chart for the index model for Provinces B after cross-validation

This gives rise to the following confusion matrix in Table 10.10 which is very similar to the one in Table 10.9 for the index model before cross-validation.

Table 10.10 Confusion Matrix of Provinces B Index Model after Cross-validation: p = 0.5

		Actual Yield	
		Low Yield	High Yield
Predicted	Low Yield	297	20
	High Yield	75	299

As for the *Time*-based logistic model, a ROC chart was produced for the index model for *Provinces B* following cross-validation. This can be used to identify the best threshold for predicting high yielding banking-halls. This ROC chart is shown in Figure 10.6.

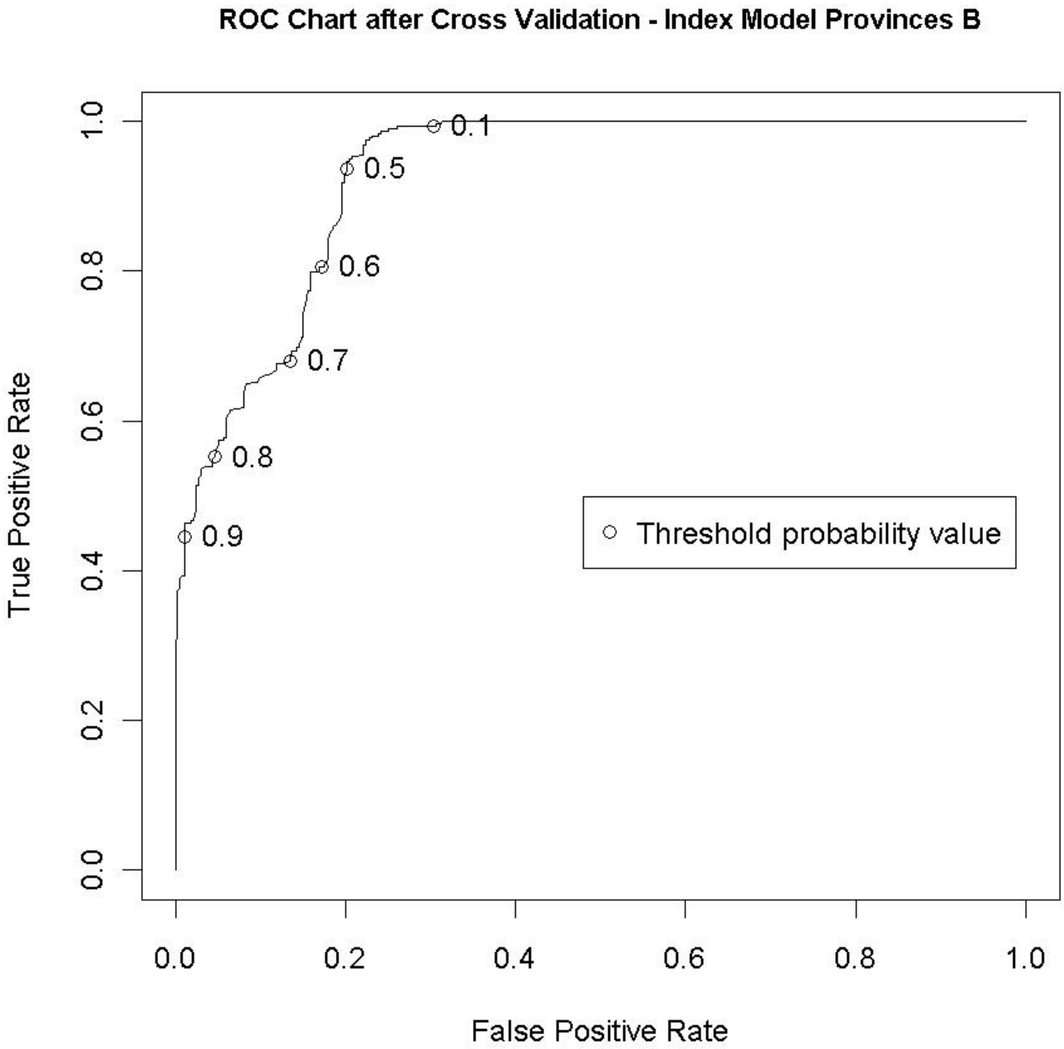


Figure 10.6 ROC Chart after Cross-Validation – Index Model *Provinces B*

Examination of the predictions bar chart and the ROC chart suggest that for an acceptably low (0.1 or less) false positive rate, a threshold of 0.8 or more is required.

A plot of the predicted positive value versus the true positive rate was also made after cross-validation. This chart is shown in figure 10.7.

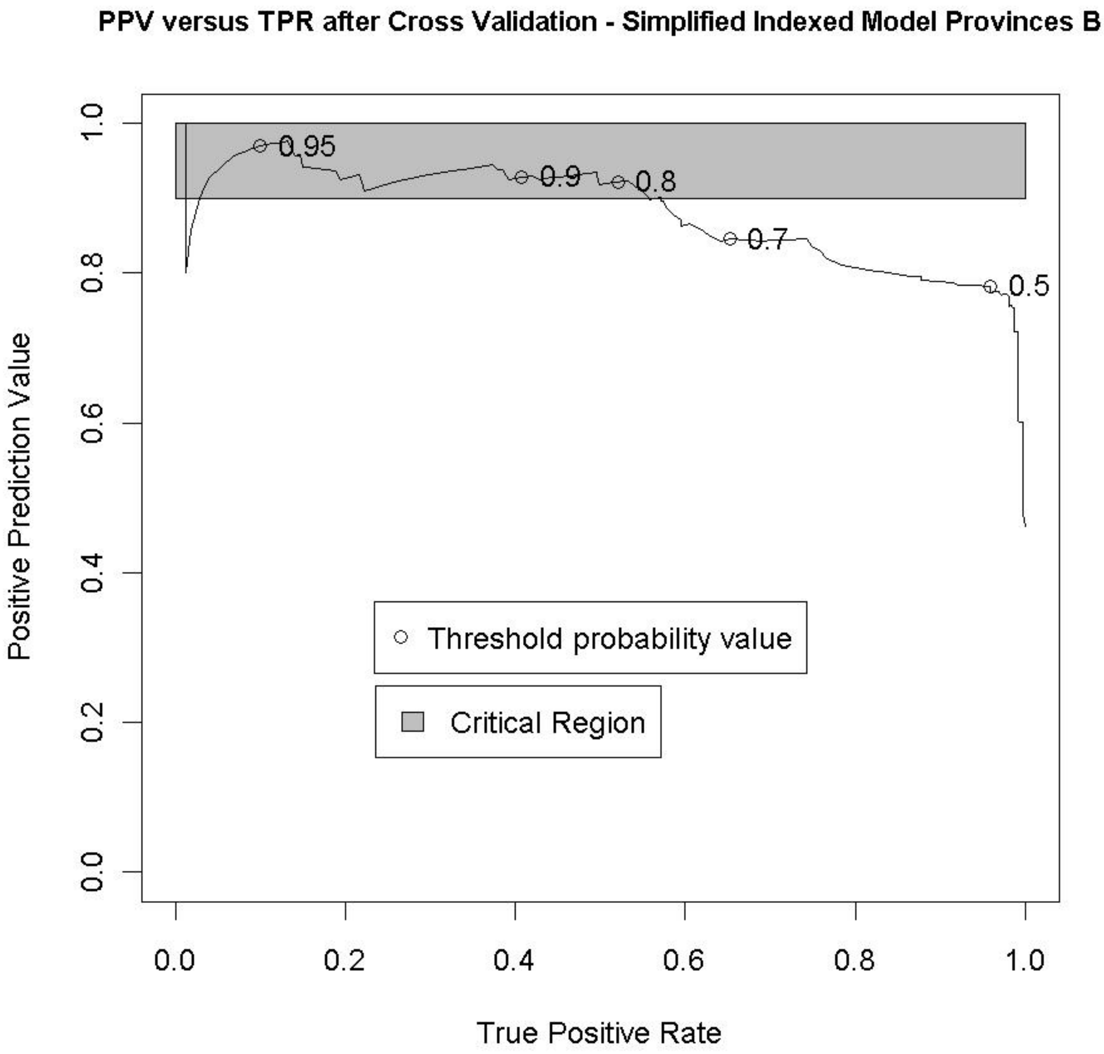


Figure 10.7 PPV versus TPR after Cross-Validation – Simplified Indexed Model *Provinces B*

The erratic behaviour of the graph in Figure 10.7 when the threshold is at an extremely high level very close to the maximum actual or predicted probability results from the very, very few actual or predicted cases which are above the threshold. The chart confirms, again, that 0.8 appears to be a good threshold.

Following the generation of the two ROC charts with respect to the basic and simplified indexed models for *Provinces B*, a further ROC chart was produced, comparing the predictive accuracy of the two models. This third ROC chart is shown in Figure 10.8. Clearly, this

ROC chart suggests that there is very little difference between the predictive powers of the basic and simplified indexed models with respect to *Provinces B*. Indeed, an eyeball examination of the chart suggests that the simplified model may actually have marginally greater predictive power than the basic model. The simplified model is, therefore, preferred.

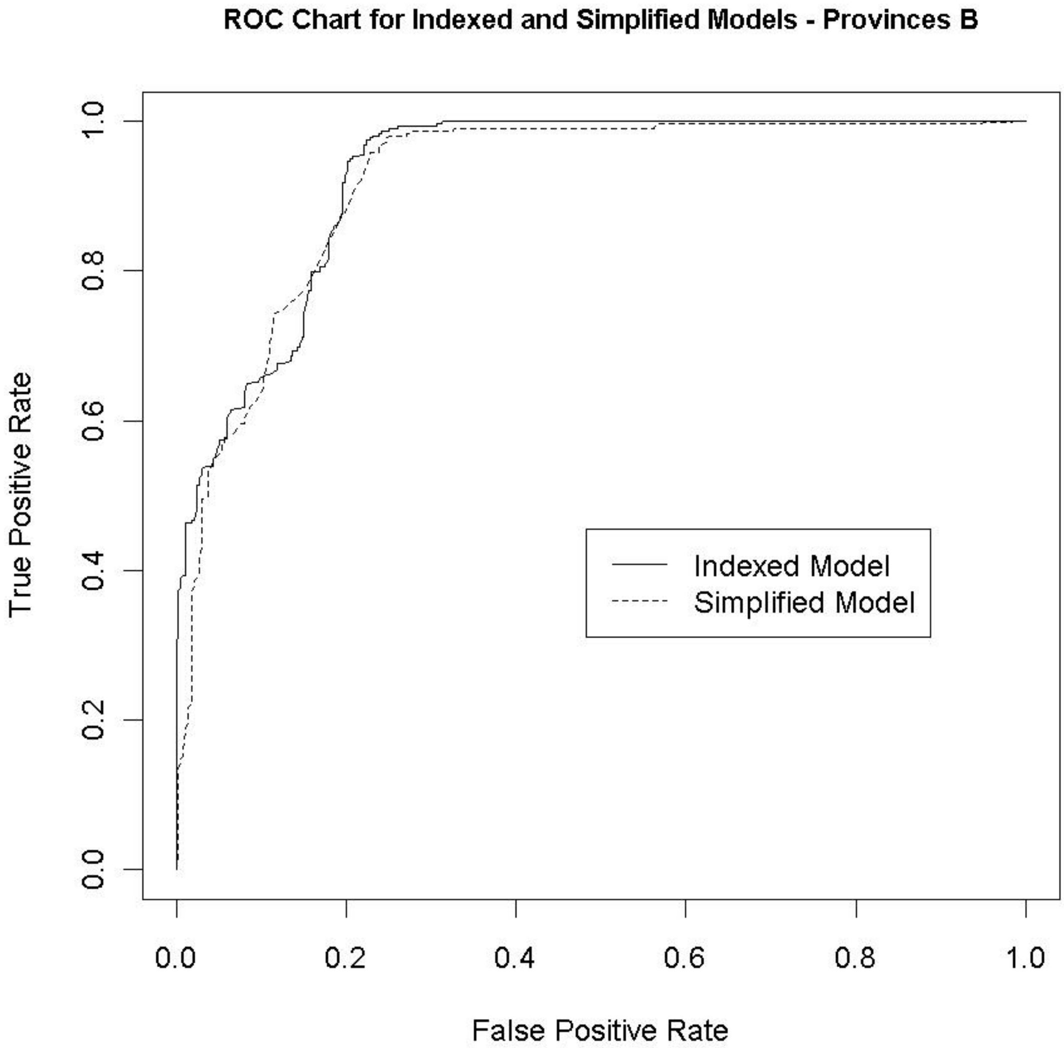


Figure 10.8 ROC Chart for Indexed and Simplified Models – *Provinces B*

10.2.4 Comparing the index models for *Provinces A* and *Provinces B*

ROC charts are also a useful way of comparing the index models built respectively from the *Provinces A* and the *Provinces B* datasets. Therefore, the index models for both datasets were plotted on the same ROC chart, which is shown in Figure 10.9. Over the region of interest, the *Provinces A* index model line is clearly closer than the *Provinces B* index model line to the top left-hand corner. Therefore, the *Provinces A* index model is shown to be the better model.

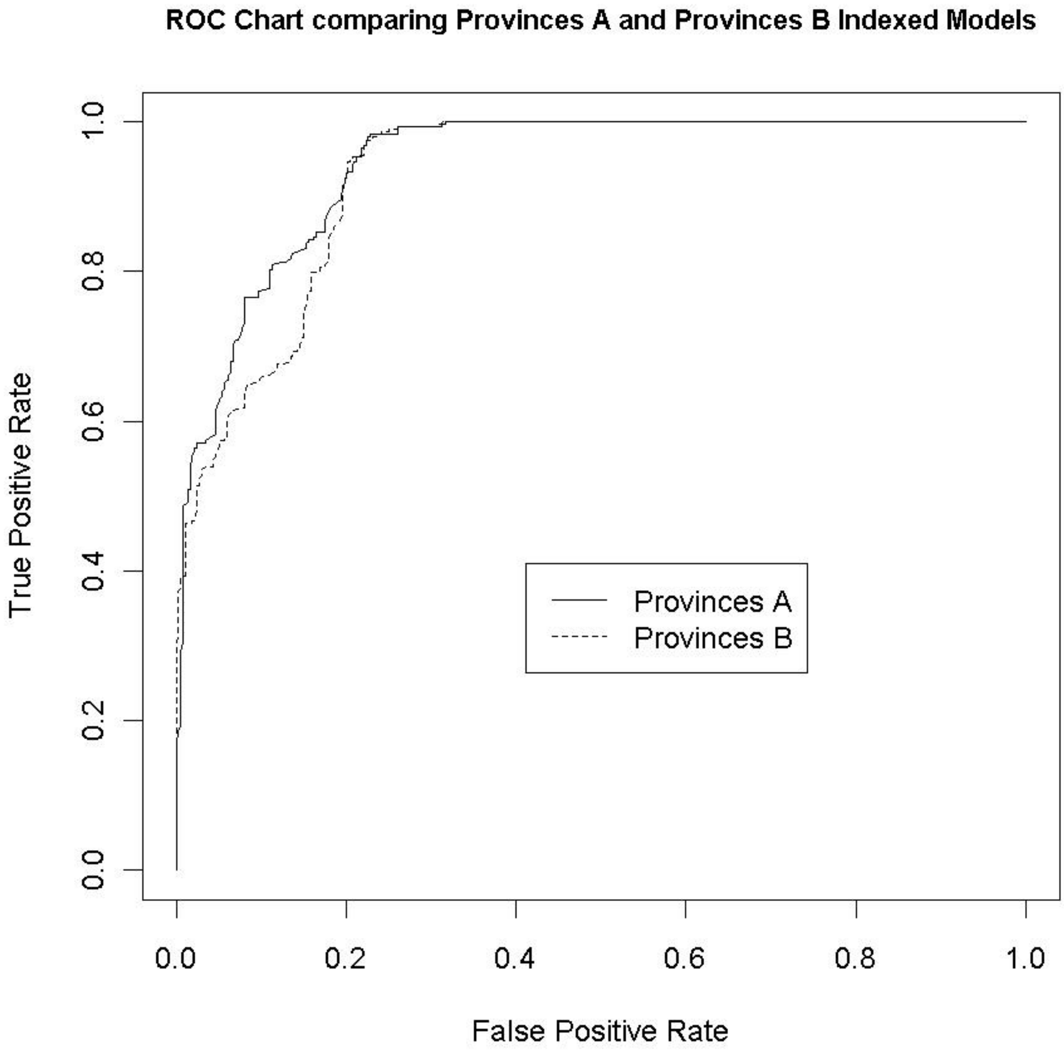


Figure 10.9 ROC Chart comparing *Provinces A* and *Provinces B* Indexed Models

10.2.5 Conclusions to logistic regression index models

For both datasets, the indexed model, using the IPD index variable, was shown to be more useful than the simplified model using the Time variable. Therefore, the index models were preferred on the grounds of utility.

The index model for both the *Provinces A* and *Provinces B* datasets comprised the same terms. Namely, the three linear terms: Lot size, Province and Index.

In an apparent contrast to the findings of the qualitative study the term *Bank* has been excluded from the model. Exclusion was on the basis of a lack of practical significance, since the term remained statistically significant. Nevertheless, the *Bank* term and the *Bank:Index* interaction were the most statistically significant terms to be omitted from the index model.

The index model for *Provinces A* is shown to have greater predictive accuracy than the index model for *Provinces B*. Therefore, the index model for Provinces A was the preferred of all the logistic regression models. For this model, a predictive threshold at above 0.9 gives especially good results. However, it would appear that at a threshold above 0.8, this model continues to provide a high positive prediction for the assembly of a banking-hall property investment portfolio. The 0.8 level is selected, because it gives an acceptable level of false positives and therefore provides a sufficient number of true positives.

10.3 ANOVA

10.3.1 Introduction to ANOVA

The same procedures were adopted for the logistic regression in this section as those that had been used when *Time*, rather than *Index*, had been an independent variable.

Whereas *Time* was a factor, or categorical variable, *Index* is a continuous variable. Therefore, the *Index* variable does not show factor levels in the equations following.

10.3.2 The Model for the *Provinces A* and *Provinces B* datasets

The full, or saturated, binary logistic model for the two datasets with dependent variable *Yield Group* (Y) has as predictor variables the three factors, *Bank*, *Lot Size* and *Provinces*, and the numeric variable, *Index*, and takes the:

$$
\begin{aligned}
Y_i = \; & \lambda && \text{(1 constant term)} \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_{I} I_i && \text{(4 first order terms)} \\
& + \lambda_{BLjk} + \lambda_{BPjl} + \dots\dots\dots + \lambda_{PIl} I_i && \text{(6 second order terms)} \\
& + \lambda_{BLPjkl} + \lambda_{BLIjl} I_i + \dots\dots\dots + \lambda_{LPIkl} I_i && \text{(4 third order terms)} \\
& + \lambda_{BLPIjkl} I_i && \text{(1 fourth order term)}
\end{aligned}
\tag{10.7}
$$

Where:

- Y_i is the predicted yield for the i'th case,
- I_i is the index value for the i'th case and
- *j, k* and *l* represent the categories of the factors *B, L,* and *P* respectively.

10.3.3 Analysis and results

10.3.3.1 Analysis and results for Provinces A

The saturated ANOVA model, comprising all the terms up to and including the fourth order, is shown in the equation 10.7. Backward elimination from the saturated model using the Akaike Information Criterion (AIC) suggested that no terms should be removed from the model. Therefore, the fourth order term was tested explicitly on an F-ratio test. The F-ratio test showed that the fourth order term was not significant at the 10 per cent level. Therefore, the fourth order term was removed from the model before backward elimination was continued.

That produced a model with six terms in the predictor variables.

$$\begin{aligned} Y_i \quad = \quad & \lambda \\ & + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_I I_i \\ & + \lambda_{BPjl} + .\lambda_{PIl} I_i \end{aligned} \tag{10.8}$$

These comprised the four first order terms, *Bank, Lot Size, Index* and *Provinces A*, and the following two second order interactive terms:

- *Bank*Index*
- *Bank*Provinces A*

The significances of the two second order interactions were then tested explicitly. The *Bank*Index* interaction is significant, but at only just beyond the ten per cent level at 9.574 per cent. The *Bank*Provinces A* term is significant at well beyond the 0.1 per cent level. Therefore, both terms were retained in the model.

Table 10.11 Analysis of Deviance Table for the Provinces A Index Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Bank	2	517.82	258.91	395.4767	< 2.2e-16	***
Index	1	416.49	416.49	636.1782	< 2.2e-16	***
Lot Size	1	5.50	5.50	8.4075	0.0038582	**
Provinces A	3	41.55	13.85	21.1578	4.149e-13	***
Bank:Index	2	4.01	2.01	3.0647	0.0473169	*
Bank:Provinces A	6	17.02	2.84	4.3330	0.0002659	***
Residuals	675	441.91	0.65			

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

In the summary of ANOVA table shown in Table 10.11, the *Bank* and *Index* terms explain much more of the variance than the other terms. The *Bank* term accounts here in the ANOVA for much more of the variance than it did in the logistic regression.

The linear terms *Bank, Index* and *Provinces A* were all found to be significant at well beyond the 0.1% level in the F-test. The linear term *Lot Size* was found to be significant at the one per cent level. The second order term *Bank:Provinces A* was found to be significant at the 0.1% level. The second order term *Bank:Index* was found to be significant at the five per cent level.

The model itself is shown by the reduced Analysis of Variance table (Table 10.12) to be significant at well beyond the 0.1% level (F is 102.08 on 15 and 675 degrees of freedom).

Table 10.12 Significance of Fitted Model for Provinces A Index Model

```
Model 1: yield ~ 1

Model  2:  yield  ~  bank  +  index  +  lotsize  +  provincesa  +
bank:index + bank:provincesa


      Resid.Df   Res.Sum of Sq  Df   Sum of Sq    F     Pr(>F)

Model 1   690      1444.32
Model 2   675       441.91     15     1002.4   102.08   < 2.2e-16
```

The model was then checked for its goodness of fit. This was done by making a plot of the residuals against the fitted values to check for systematic changes in variance. The residuals are the difference between the actual yield and the fitted yield. Plotting the residuals against the fitted yield will mainly show if there is any trend in the variance of the residuals as the fitted value increases.

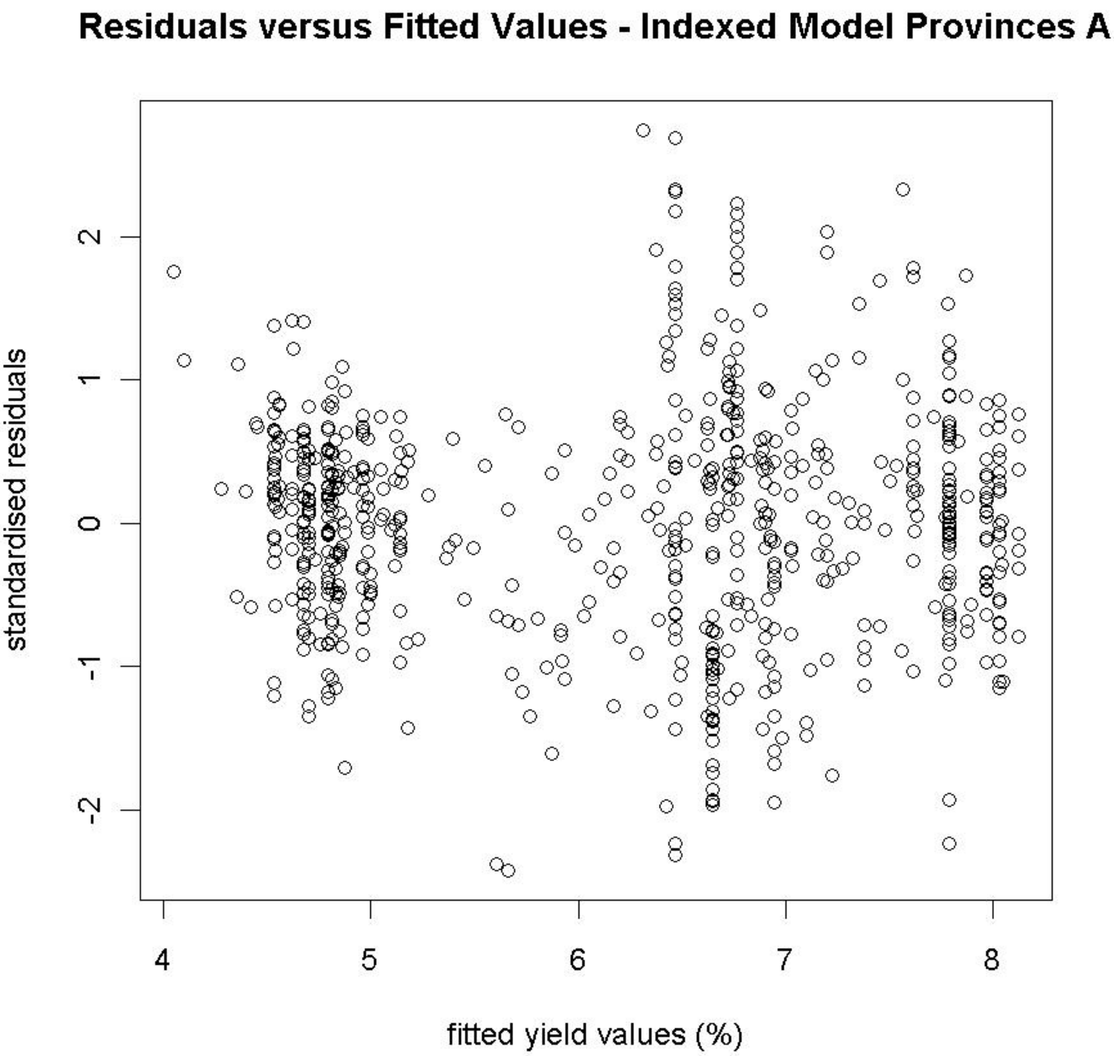


Figure 10.10 Residuals versus Fitted Values – Indexed Model *Provinces A*

A visual examination of the residual versus fitted values plot (Figure 10.10) suggested that there were no clear systematic changes in variance. As far as the ANOVAs are concerned, the commonly used variance stabilizing transformations (e.g. square root and log) are unlikely to be effective in this case.

A Q-Q plot of the residuals (Figure 10.11) was made to give a visual check for normality. The tail ends are off the fitted line, which suggests that the distribution has a greater peak than a normal distribution would have. This plot, in conjunction with the kurtosis value, shows that the distribution is leptokurtic. ANOVA tests are known to be particularly sensitive to this type of departure from normality (Wetherill, 1981: 20).

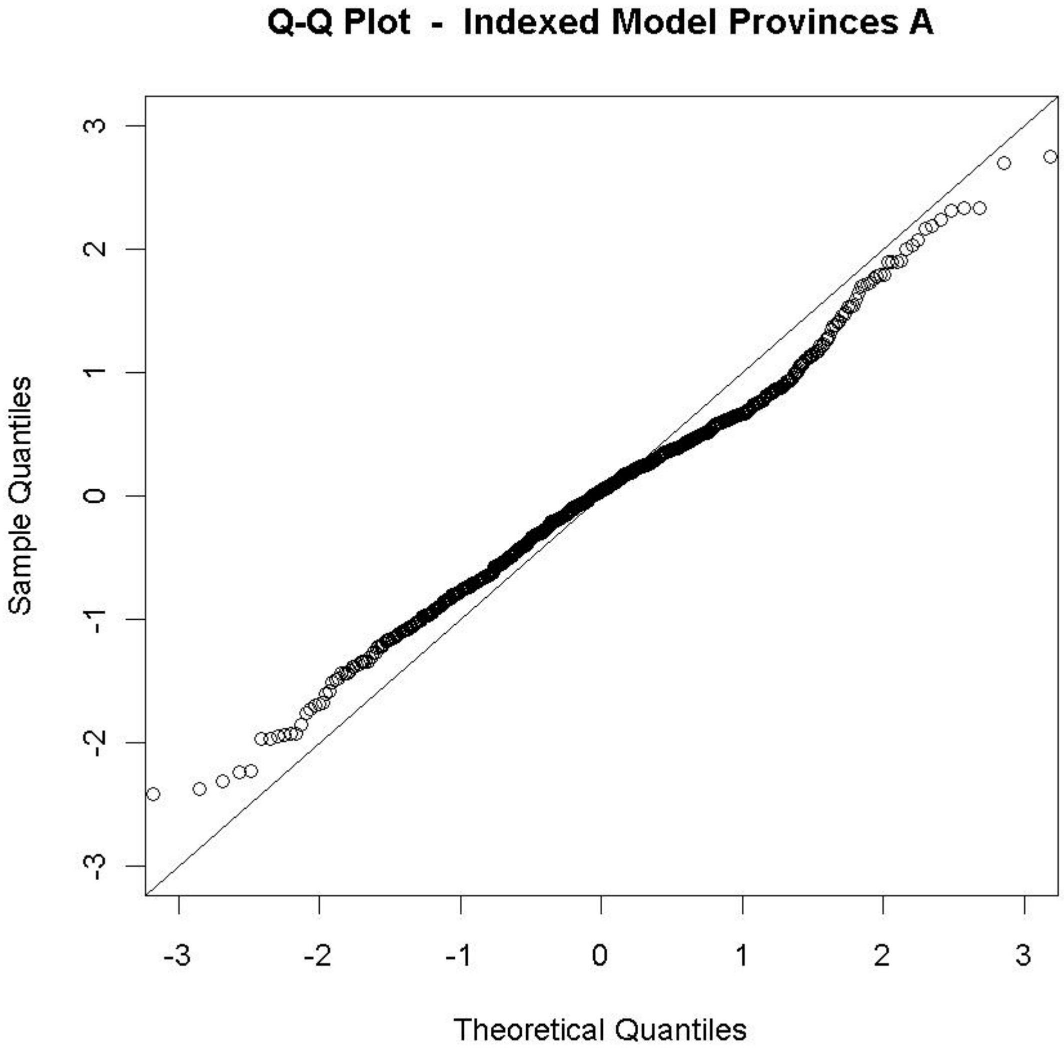


Figure 10.11 Q-Q Plot – Indexed Model *Provinces A*

A confusion matrix was derived from the plot (similar to that in Figure 10.12) of Predicted Yield against Actual Yield. A threshold of 6.35 per cent was applied to both yield values. This divided the plot into four quadrants from which a confusion matrix could be generated. The confusion matrix is reproduced in Table 10.13.

Table 10.13 Confusion Matrix of Provinces A ANOVA Model before cross-validation: Threshold value of 6.35%

		Predicted	
		Low Yield	High Yield
Actual	Low Yield	291	82
	High Yield	12	306

Cross-validation was performed using random splits into ten groups without replication. The identical confusion matrix was produced after cross-validation and is reproduced in Table 10.14. As in the logistic regression, the cross-validation results for the ANOVA are in general no worse than before cross-validation.

Table 10.14 Confusion Matrix of Provinces A ANOVA Model after Cross-validation: Threshold value of 6.35%

		Predicted	
		Low Yield	High Yield
Actual	Low Yield	291	82
	High Yield	12	306

In order to look at the practical usefulness of the results, the predicted yield was plotted against the actual yield. This plot is shown in Figure 10.12. The horizontal and vertical lines in the figure represent a threshold of 6.35 per cent. The confusion matrix is derived from the number of points in each quadrant. There is a closer relationship between predicted and actual yields using *Index* as a variable than was using *Time* as a variable.

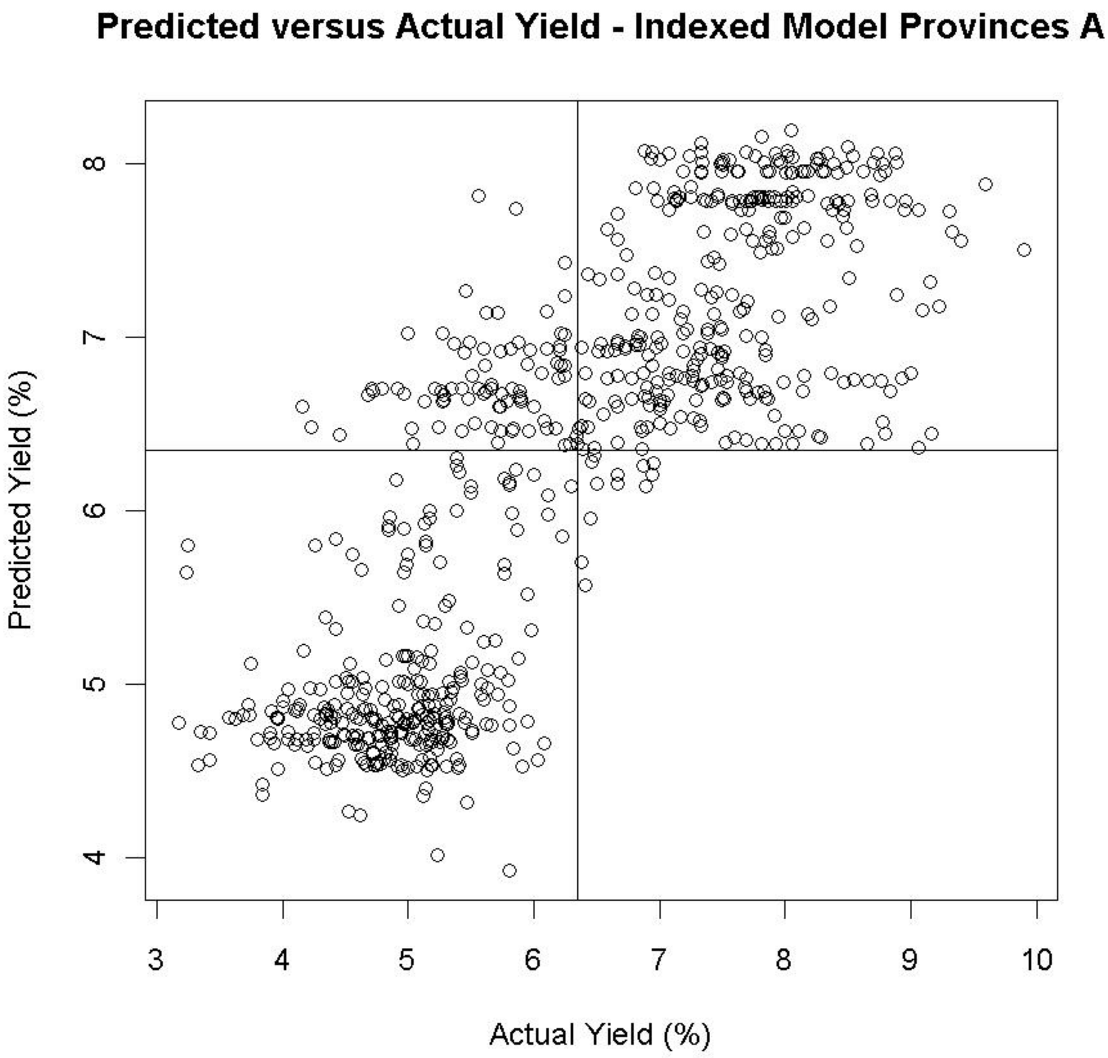


Figure 10.12 Predicted versus Actual Yield – Indexed Model *Provinces A*

By looking at the effects of applying different thresholds to Figure 10.12, a ROC chart similar to those in the logistic analyses can be produced.

Figure 10.13 shows the Predicted Positive Value versus True Positive Rate after cross-validation with respect to the indexed ANOVA model for *Provinces A*.

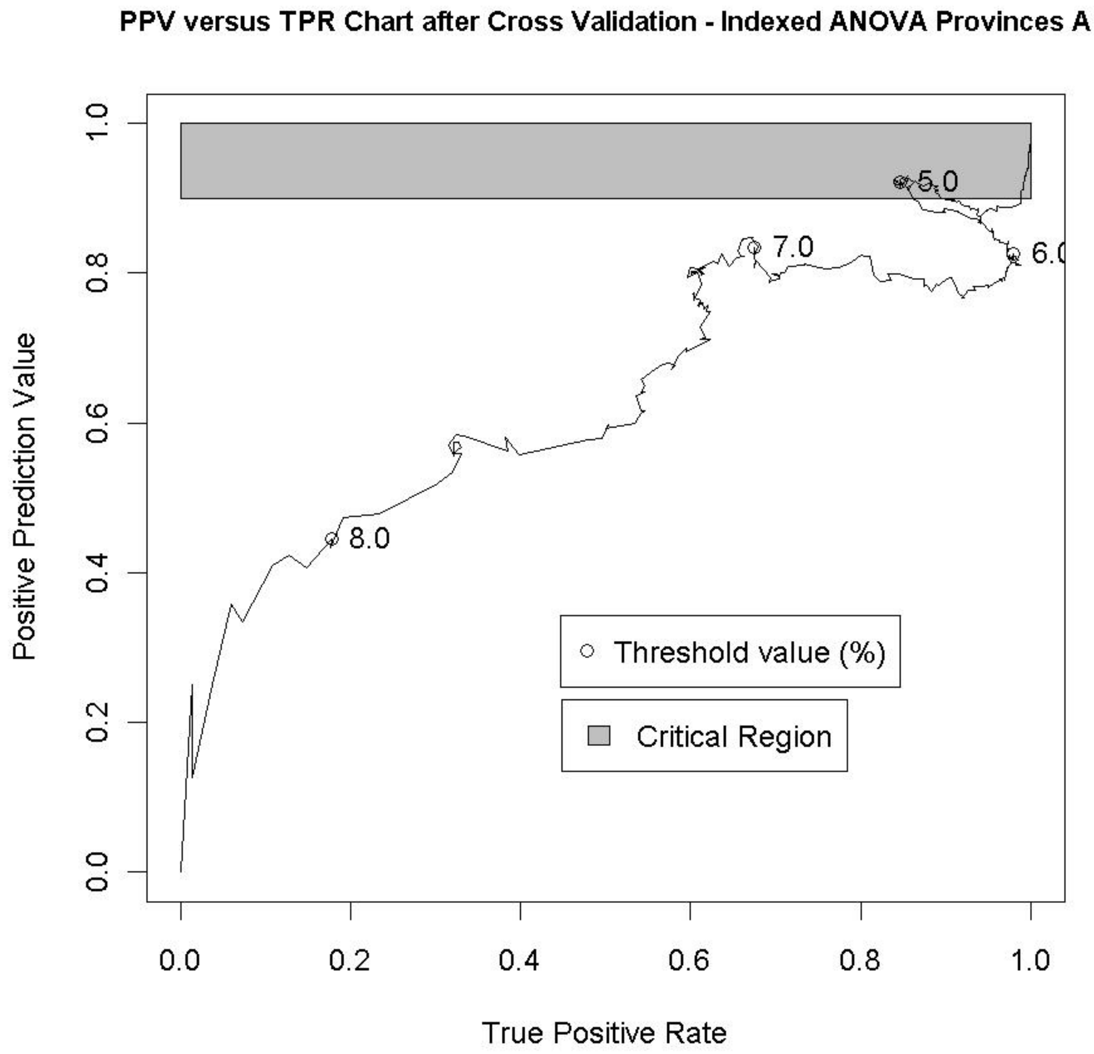


Figure 10.13 PPV versus TPR Chart after Cross-Validation - Indexed ANOVA *Provinces A*

The Predictive Positive Value versus True Positive Rate chart in Figure 10.13 is very erratic with the line of the plots looping in several places. Moreover, apart from a very small area within one of the loops, none of the selected predictive positive rate of ninety per cent or more for high yield was achieved.

Figure 10.13 can be compared and contrasted with Figure 10.3 (logistic regression) which also shows a Predicted Positive Value versus True Positive Rate and which chart confirms that 0.8 appears to be a good threshold.

Figure 10.13 demonstrates that the ANOVA model for *Provinces A* does not provide a meaningful basis for selecting the threshold required in order to make useful predictions for investing in retail bank premises.

10.3.3.2 Analysis and results of Provinces B

The saturated ANOVA model, comprising all the terms up to and including the fourth order, was shown in the equation 10.7. Backward elimination from the saturated model using the Akaike Information Criterion (AIC) suggested that no terms should be removed from the model. Therefore, the fourth order term was tested explicitly in an F-ratio test. The F-ratio test showed that the fourth order term was not significant at the 10 per cent level. Therefore, the fourth order term was removed from the model before backward elimination was continued.

That produced the same model for *Provinces B* as that for *Provinces A*. Thus, it produced a model with six terms in the predictor variables as expressed in the following equation:

$$
\begin{aligned}
Y_i \quad = \quad & \lambda \\
& + \lambda_{Bj} + \lambda_{Lk} + \lambda_{Pl} + \lambda_I I_i \\
& + \lambda_{BPjl} + .\lambda_{PIl} I_i
\end{aligned} \tag{10.9}
$$

These comprised the four first order terms, *Bank, Lot Size, Index* and *Provinces B*, and the following two second order interactive terms:

- *Bank*Index*
- *Bank*Provinces B*

The significance of the two second order interactions were then tested explicitly. The *Bank*Index* interaction is significant at beyond the five per cent level at 1.753 per cent when explicitly tested. The *Bank*Provinces B* term is significant at the one per cent level at 0.1352 per cent when explicitly tested. Therefore, both terms were retained in the model.

Table 10.15 Analysis of Deviance Table for the Provinces B Index Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Bank	2	517.82	258.91	378.0291	< 2.2e-16	***
Index	1	416.49	416.49	608.1114	< 2.2e-16	***
Lot Size	1	5.50	5.50	8.0366	0.004721	**
Provinces B	3	22.16	7.39	10.7832	6.288e-07	***
Bank:Index	2	5.15	2.57	3.7591	0.023794	*
Bank:Provinces B	6	14.89	2.48	3.6226	0.001518	**
Residuals	675	462.31	0.68			

Significance codes 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Terms in the table added sequentially (first to last). In this table the significances are dependent on the order in which the terms are inserted. The true significance of a term is determined by examining the effect on the model when that term, and that term alone, is removed.

In the summary of ANOVA table shown in Table 10.15, the *Bank* and *Index* terms explain much more of the variance than the other terms. The *Bank* term accounts here in the ANOVA for much more of the variance than it did in the logistic regression.

The linear terms *Bank, Index* and *Provinces B* were all found to be significant at well beyond the 0.1% level in the F-test. The linear term *Lot Size* was found to be significant at the one per cent level. The second order term *Bank:Provinces B* was found to be significant at the one per cent level. The second order term *Bank:Index* was found to be significant at the five per cent level.

The model itself is shown by the reduced Analysis of Variance table (Table 10.16) to be significant at well beyond the 0.1% level (F is 95.587 on 15 and 675 degrees of freedom).

Table 10.16 Significance of Fitted Model for Provinces B Index Model

```
Model 1: yield ~ 1

Model 2: yield ~ bank + index + lotsize + provincesb +
bank:index + bank:provincesb


      Resid.Df  Res.Sum of Sq  Df  Sum of Sq      F     Pr(>F)

Model 1   690      1444.32
Model 2   675       462.31     15   982.01     95.587  < 2.2e-16
```

The model was then checked for its goodness of fit. Again this was done by making a plot of the residuals against the fitted values to check for systematic changes in variance.

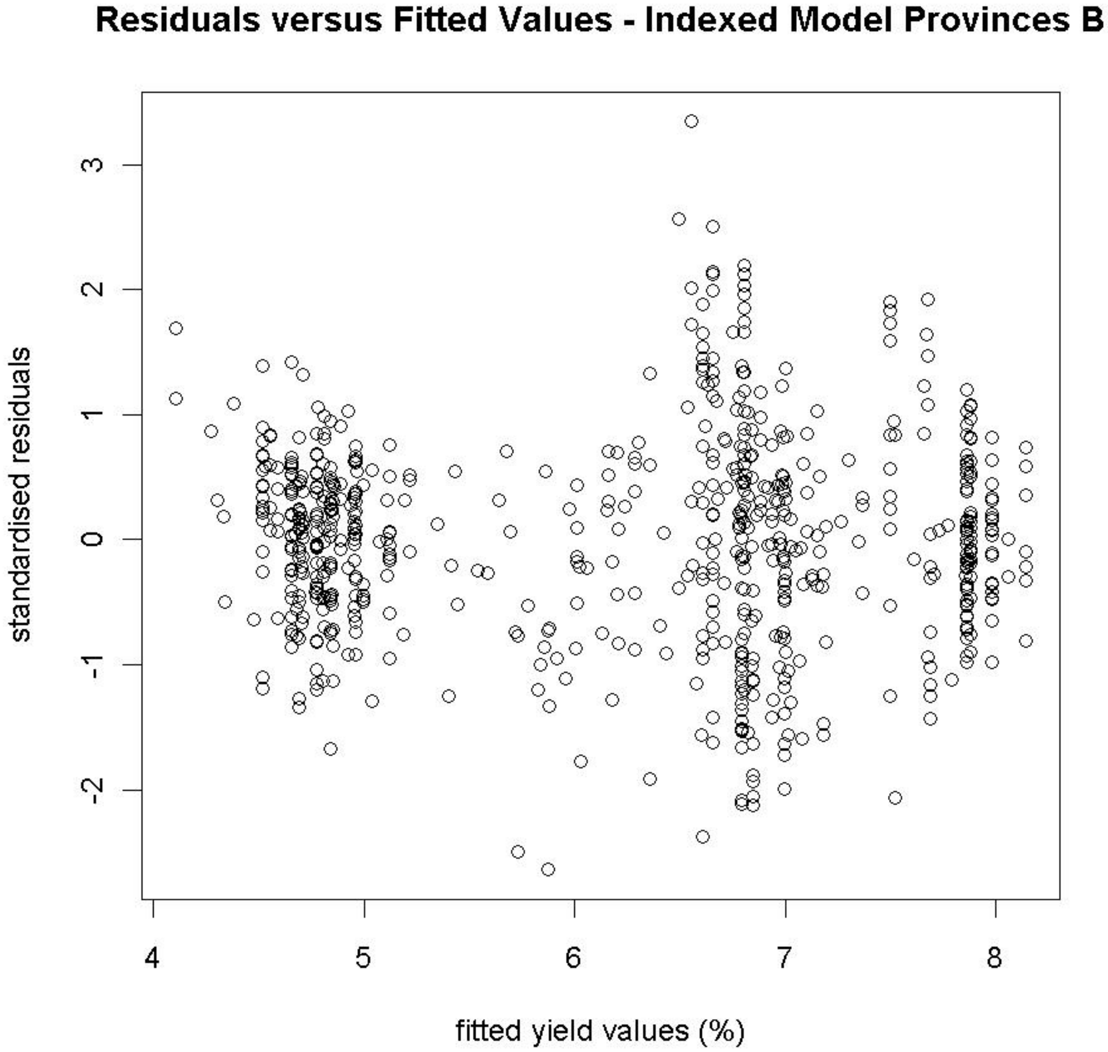


Figure 10.14 Residuals versus Fitted Values – Indexed Model *Provinces B*

A visual examination of the residual versus fitted values plot (Figure 10.14) suggested that there were no clear systematic changes in variance. Again, as far as the ANOVAs are concerned, the commonly used variance stabilizing transformations (e.g. square root and log) are unlikely to be effective in this case.

A Q-Q plot of the residuals (Figure 10.15) was made to give a visual check for normality. The tail ends are off the fitted line, which suggests that the distribution has a greater peak than a normal distribution would have. This plot, in conjunction with the kurtosis value, shows that the distribution is leptokurtic. ANOVA tests are known to be particularly sensitive to this type of departure from normality.

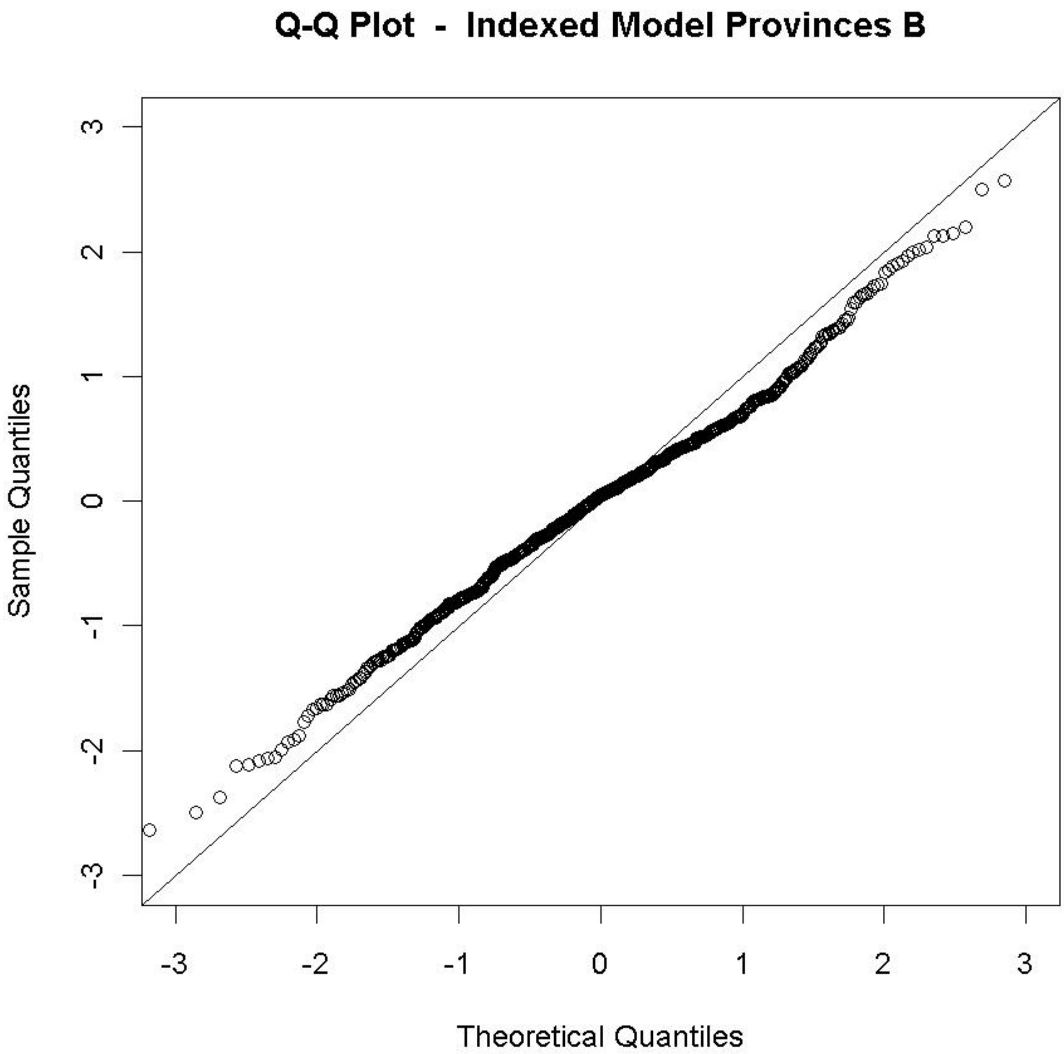


Figure 10.15 Q-Q Plot – Indexed Model *Provinces B*

A confusion matrix was derived from the plot (similar to that in Figure 10.16) of Predicted Yield against Actual Yield. A threshold of 6.35 per cent was applied to both yield values. This divided the plot into four quadrants from which a confusion matrix could be generated. The confusion matrix is reproduced in Table 10.17.

Table 10.17 Confusion Matrix of Provinces B ANOVA Model before cross-validation: Threshold value of 6.35%

		Predicted	
		Low Yield	High Yield
Actual	Low Yield	292	81
	High Yield	13	305

Cross-validation was performed using random splits into ten groups without replication. The identical confusion matrix was produced after cross-validation and is reproduced in Table 10.18. As in the logistic regression, the cross-validation results for the ANOVA are in general no worse than before cross-validation. The strong similarity of the confusion matrices before and after cross-validation shows that the model validity is good.

Table 10.18 Confusion Matrix of Provinces B ANOVA Model after Cross-validation: Threshold value of 6.35%

		Predicted	
		Low Yield	High Yield
Actual	Low Yield	292	81
	High Yield	15	303

In order to look at the practical usefulness of the results, the predicted yield was plotted against the actual yield. This plot is shown in Figure 10.16. The horizontal and vertical lines in the figure represent a threshold of 6.35 per cent. The confusion matrix is derived from the number of points in each quadrant. There is a closer relationship between predicted and actual yields using *Index* as a variable than was using *Time* as a variable.

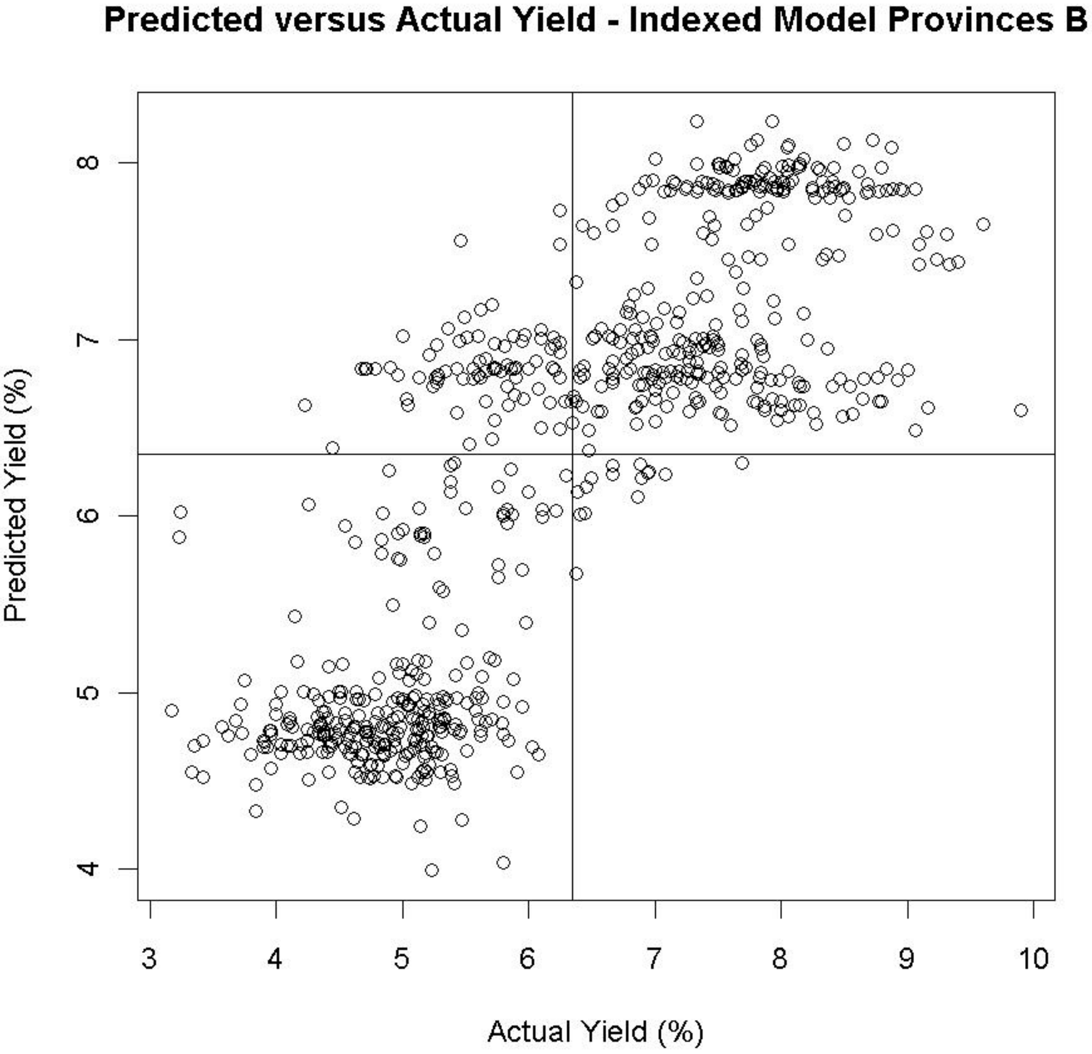


Figure 10.16 Predicted versus Actual Yield – Indexed Model *Provinces B*

By looking at the effects of applying different thresholds to Figure 10.16, a ROC chart similar to those in the logistic analyses can be produced.

Figure 10.17 shows the Predicted Positive Value versus True Positive Rate after cross-validation with respect to the indexed ANOVA model for *Provinces B*.

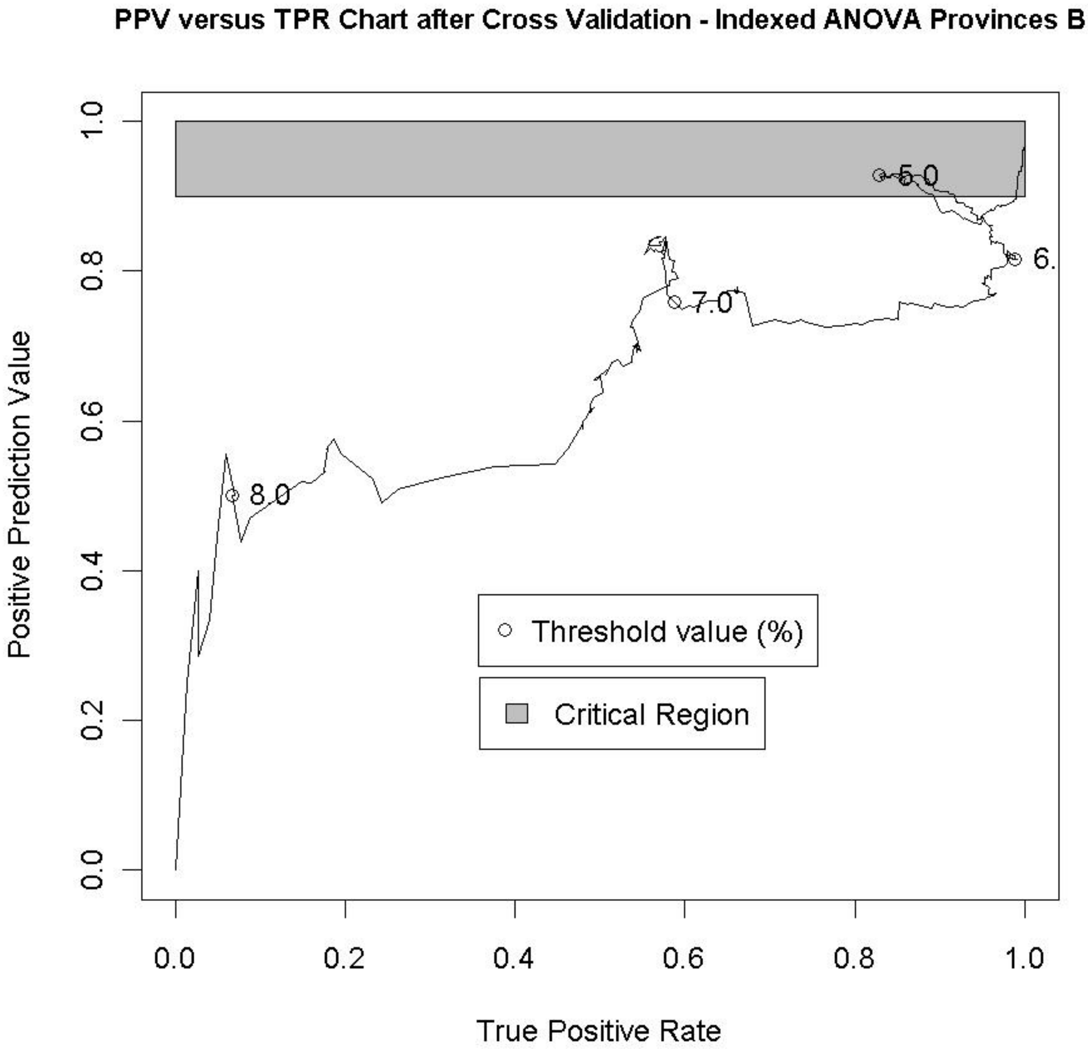


Figure 10.17 PPV versus TPR Chart after Cross-Validation – Indexed ANOVA *Provinces B*

The Predictive Positive Value versus True Positive Rate chart in Figure 10.17 is very erratic with the line of the plots looping in several places. Moreover, apart from a very small area within one of the loops, none of the selected predictive positive rate of ninety per cent or more for high yield was achieved.

Figure 10.17 can be compared and contrasted with Figure 10.7 (logistic regression) which also shows a Predicted Positive Value versus True Positive Rate and which chart confirms that 0.8 appears to be a good threshold.

Figure 10.17 demonstrates that the ANOVA model for *Provinces B* does not provide a meaningful basis for selecting the threshold required in order to make useful predictions for investing in retail bank premises.

10.3.4. Comparison between Provinces A and Provinces B Indexed ANOVAs

For the sake of completeness, Figure 10.18 compares the ROC charts for both the Provinces A and the Provinces B indexed ANOVA models. Notwithstanding that each of these two models has been identified as not being useful, Figure 10.18 does not show one as being any better than the other. One is no better than the other, because the line for each crosses the other repeatedly. Therefore, there is no indication that one is to be preferred to the other.

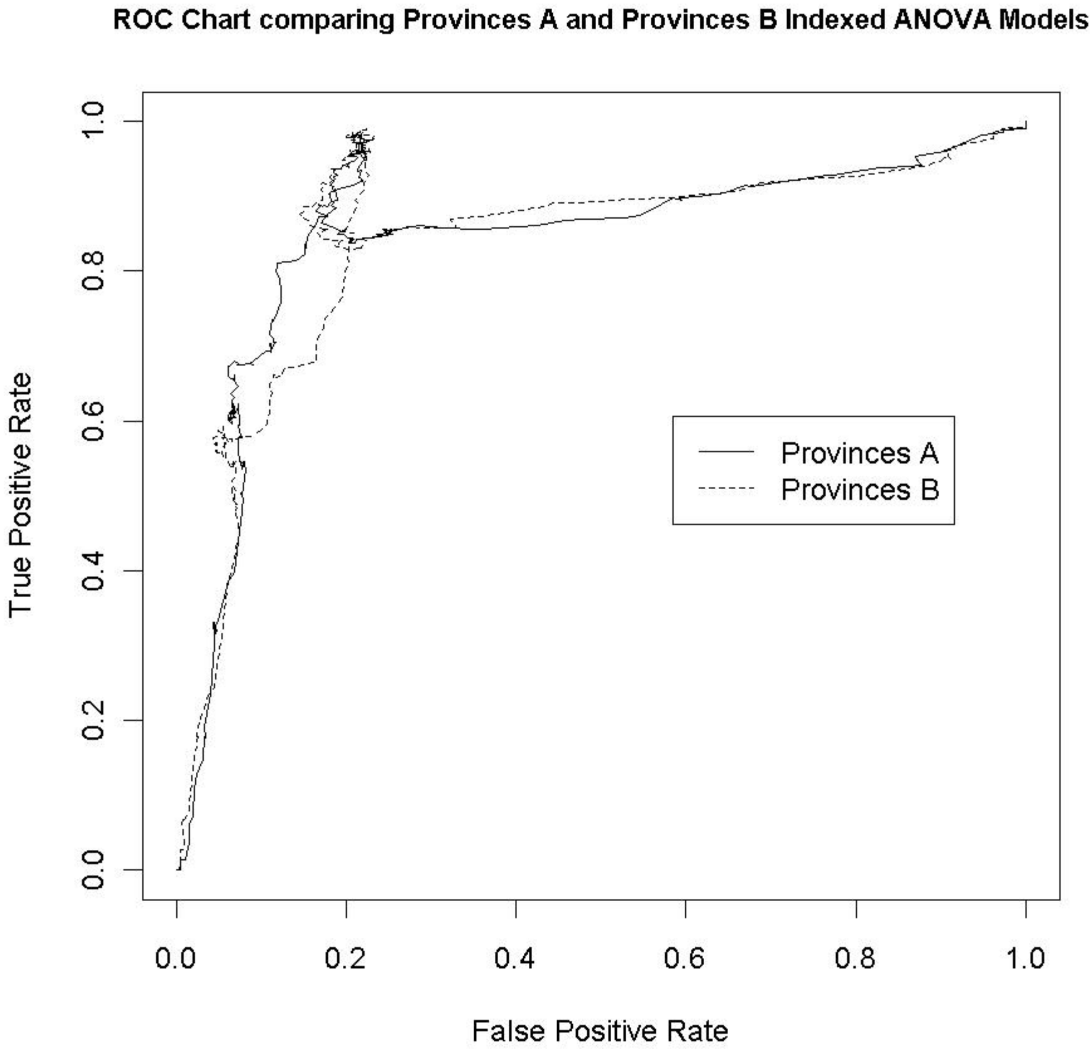


Figure 10.18 ROC Chart comparing *Provinces A* and *Provinces B* Indexed ANOVA Models

10.3.5. Comparison between Timed and Indexed ANOVA models

For the sake of completeness, Figures 10.19 and 10.20 respectively compare *Provinces A* and *Provinces B* ROC charts for the time and indexed ANOVA models. Notwithstanding that each of the two models for the respective datasets has been identified as not being useful, Figures 10.19 and 10.20 do not show one as being any better than the other. One is no better than the other, because the line for each crosses the other repeatedly. Therefore, there is no indication that one is to be preferred to the other.

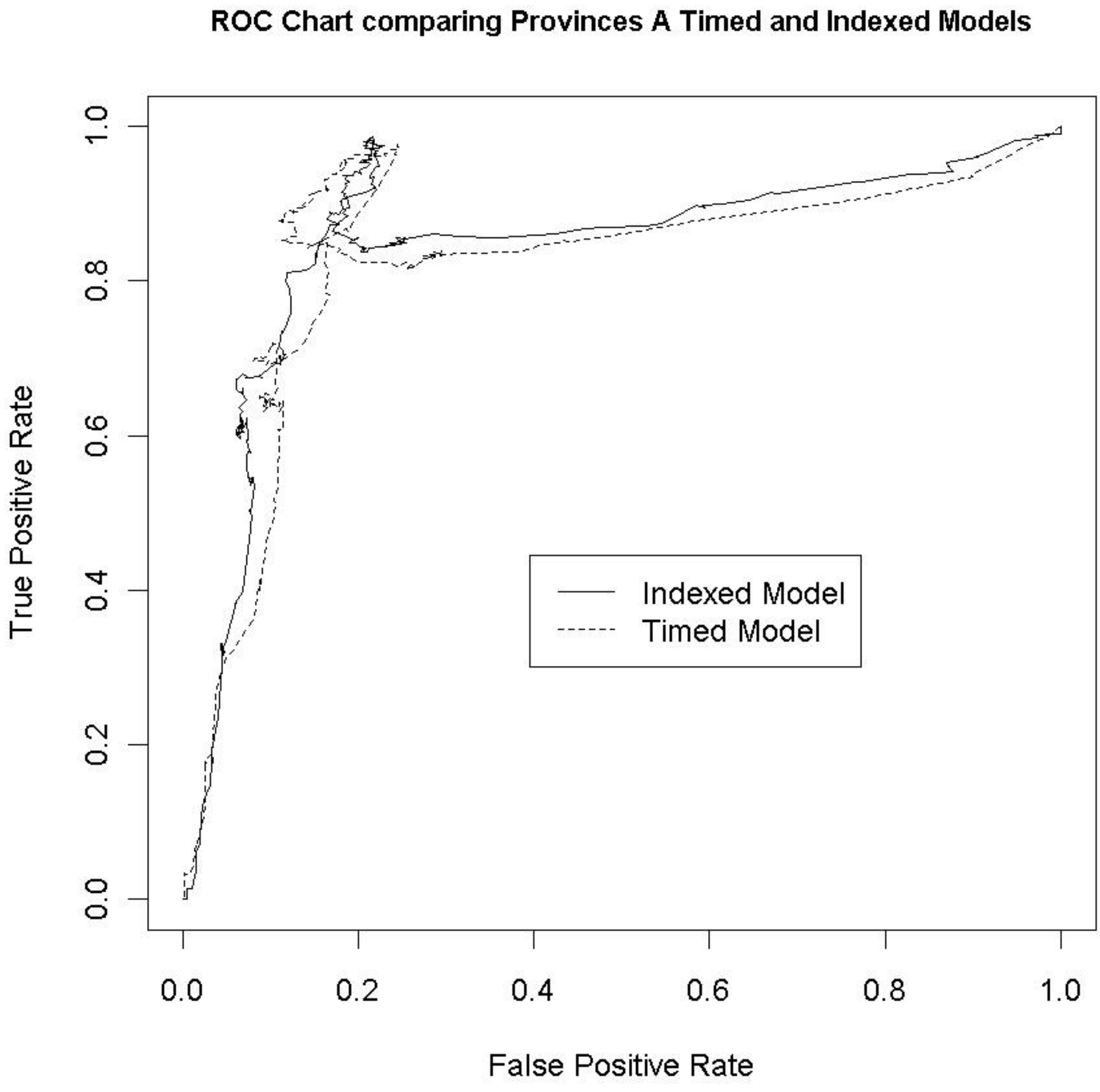


Figure 10.19 ROC Chart comparing *Provinces A* Timed and Indexed Models

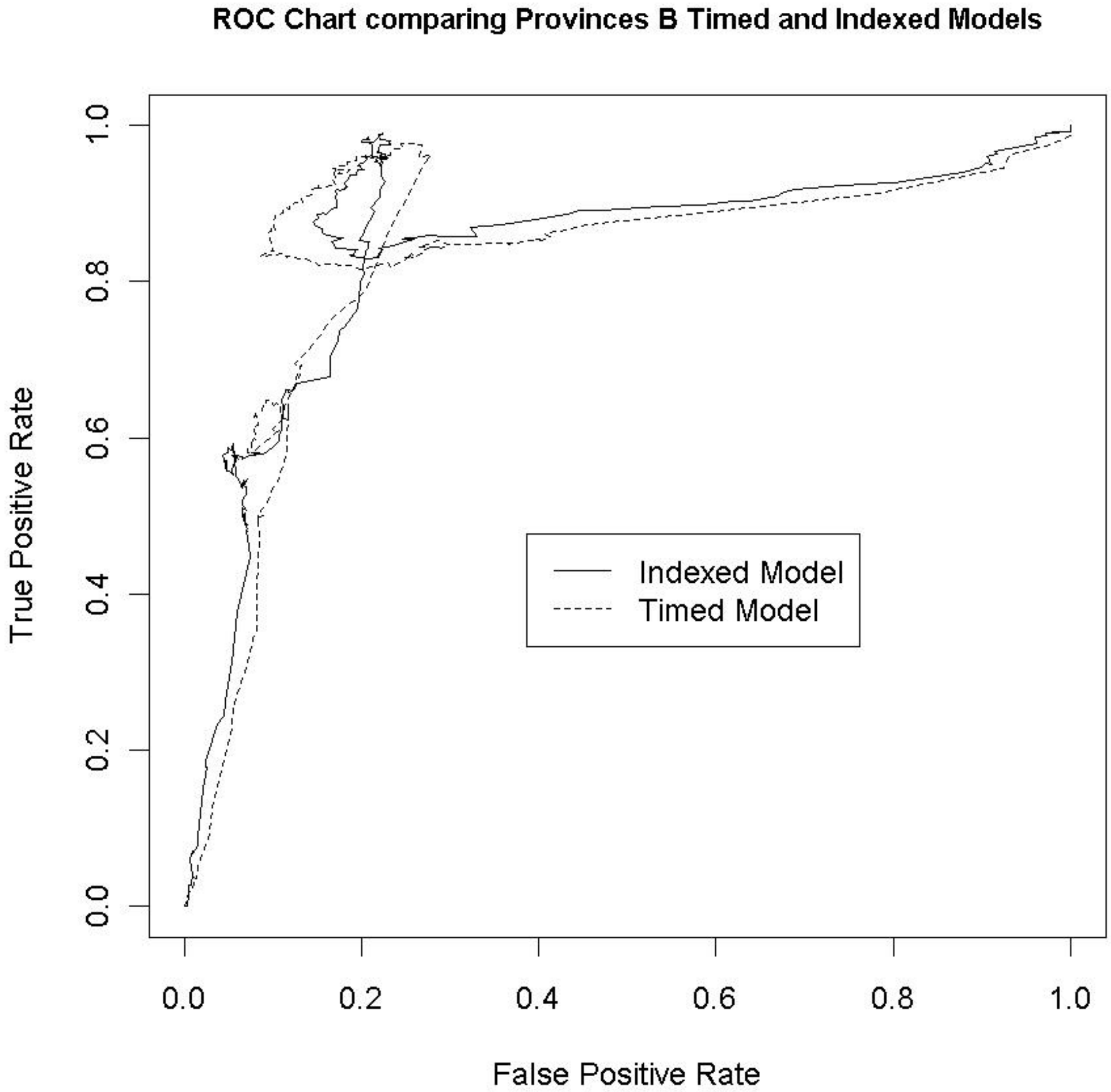


Figure 10.20 ROC Chart comparing *Provinces B* Timed and Indexed Models

10.4 Conclusions to Advanced Analyses using index data

10.4.1 Logistic Regression

- For both *Provinces A* and *Provinces B* the logistic regression analysis produced an index model comprising the same three terms in the predictor variables. These were the following three linear terms: *Lot Size, Index and Provinces*.

- In an apparent contrast to the findings of the qualitative study, the term *Bank* was excluded from the model, but this was on the grounds of practical significance rather than statistical significance. This exclusion was also in contrast to that part of the quantitative study using *Time* instead of *Index* as an independent variable.

- In practical terms for the *Provinces A* dataset, the indexed model, using the *Index* variable, was shown to be more useful than the simplified model using the *Time* variable. However, this was less clear with respect to the *Provinces B* dataset.

- In both cases, a probability threshold at or above 0.8 was found to give sufficiently accurate predictions for application for investment decisions in portfolio-building.

10.4.2 ANOVA

- For both the *Provinces A* and the *Provinces B* datasets, the ANOVA produced identical terms in the predictor variables. These were the four first order terms plus the two second order interactions *Bank*Index* and *Bank*Provinces.*

- In both cases, the distribution of predicted versus actual yield was such that the problem of selecting a suitable threshold as a basis for investment decision-making proved to be intractable. The results were erratic and no feasible criteria for a good threshold could be found.

- No useful model could be found.

10.4.3 Final observations

The form of the models in terms of the variables is the same for both the *Provinces A* and *Provinces B* versions of data collapse. This is true of both the logistic regression and ANOVA models comprising the *Index* variable instead of the *Time* one. Notwithstanding this, the values of the coefficients are not the same between the models.

For property investment decision making with respect to British banking-halls, binary logistic regression proved to produce a more useful model than ANOVA. Indeed, the ANOVA models proved to be of no utility. Since the *Provinces A* logistic regression model was found to give slightly better predictions than that for *Provinces B*, the former was preferred.

Furthermore, the *Provinces A* index model was found to be of more use than the *Provinces A* simplified model. Therefore, the *Provinces A* index model is the preferred model.

The selected model is shown to be an effective way of answering the research question, "How can property investors select freehold British banking-halls that are likely to produce the highest yield on their investment?" (see section 1.4). This conclusion is supported by the literature review (see section 4.10) and the qualitative study (see sections 6.8 and 6.9).

11 Further validation from post-study period data

11.1 Introduction

Subsequent to the observation and analyses of the data covering the study period, additional data for the eighteen months following that period were made available. Data appertaining to those banking-halls sold at auction as investment properties continued to be published in print and on the internet shortly after they were sold. Data relating to the Investment Property Databank (IPD) United Kingdom Retail Index was only made available to this research much later. These IPD data cover the period from October, 2006 until December, 2007 inclusive. In order to test the modelling further and to make it more robust, it was further validated using these additional IPD data.

The new data relating to the eighteen months following the study period comprised one hundred and ten cases. They included lots recorded on the same criteria as those included in the main study.

During the eighteen months, two banking-halls in Northern Ireland were sold. However, these were omitted from the new dataset for two reasons. Firstly, no banking-hall premises located in Northern Ireland were observed as having been sold during the study period. Secondly, the two that were sold during the eighteen months following the study period were let to banks other than those three with the largest counts of lots sold.

11.2 The validation process

11.2.1 The logistic regression model

The quantitative study had identified the *Provinces A* logistic regression indexed model as being the preferred model. Therefore, this model was subjected to further validation using the post-study period data. The further validation showed that the model remained a good one insofar as it correctly predicted 107 of the 110 cases. Since nearly all the cases belonged to the low yield group, only three lots had a true yield at or greater than 6.35 per cent. The model predicted all 110 cases as belonging to the low yield group. Of the three high yield cases, the actual yields were all at 6.6 per cent or less and therefore at the lower end of the range covered by the high yield group.

Set at the 6.35 per cent threshold, the original indexed logistic model correctly predicted 107 out of 110 cases. However, although the model remained valid, the results, that is to say the predictions of high yield group membership, were of no practical utility for investment purposes in this period. The reason for this is that yield compression arising in the overall property investment market at this time meant that the yields of nearly all of the properties were below the threshold between low and high yield. The probability bar chart in Figure 11.1 shows the low probability of the lots falling within the high yield group when the threshold for high yield is set at a yield of 6.35 per cent.

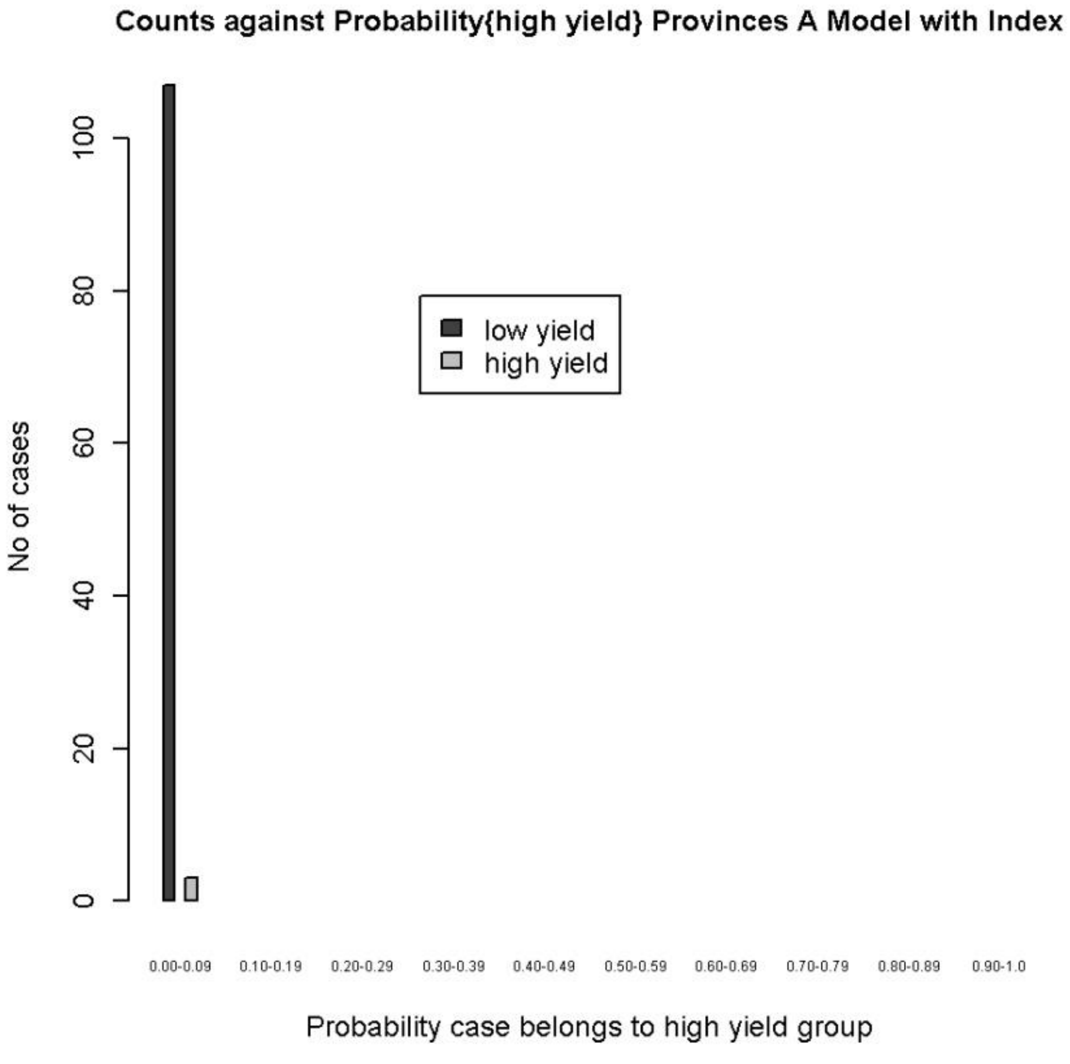


Figure 11.1 Probability bar chart for premises in the new dataset when high yield is 6.35% or more

Alternative approaches to selecting properties for investment were therefore sought. Two approaches based upon the *Provinces* A indexed logistic regression model were considered:

i Properties were ranked in order of their predicted probabilities of being in the high yield group and the ten per cent properties with the highest probability were selected for investment purposes. These are expressed in the histogram in Figure 11.2.

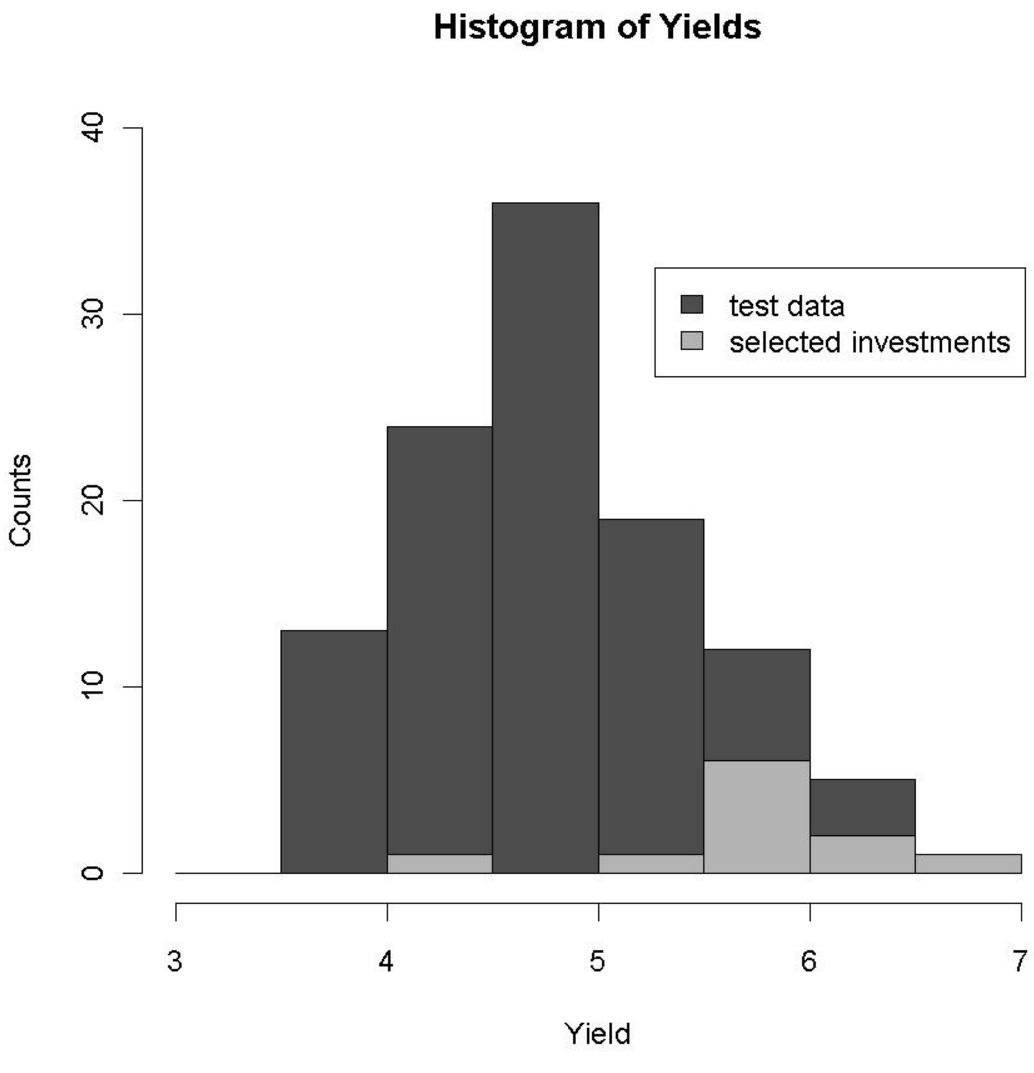


Figure 11.2 Histogram of yields for premises in new dataset when high yield is 6.35% or more

The mean yield for the selected ten per cent of properties with the highest probability of being in the high yield group was 5.73 per cent. The mean yield for all the properties in the dataset was 4.82 per cent.

ii The initial data were re-analyzed with a lower threshold between the two yield groups set at 5.5 per cent. A new logistic regression indexed model for *Provinces A* was developed using these data. This model was then applied to the new dataset. At the 5.5 per cent threshold for high yield, the model comprised all four of the linear terms plus all the second order terms with the exception of the *Bank:Provinces A* interaction. The model also comprised two third order terms. Accordingly, at the 5.5 per cent threshold the logistic regression model comprises the terms: *Bank, Lot Size, Index, Provinces A, Bank:Lot Size, Bank:Index, Index:Lot Size, Index:Provinces A, Lot Size:Provinces A, Bank:Index:Lot Size* and *Index:Lot Size:Provinces A*.

Set at the 5.5 per cent threshold, this model correctly predicted 92 out of 110 cases. Furthermore, this model produced exactly the same predictions of high and low yield groups as the original model. That is to say that all 110 cases were predicted to be in the low yield group even when the threshold on predicted probability was set at 0.5. The probability bar chart in Figure 11.3 shows the low probability of the lots falling within the high yield group when the threshold for high yield is set at a yield of 5.5 per cent.

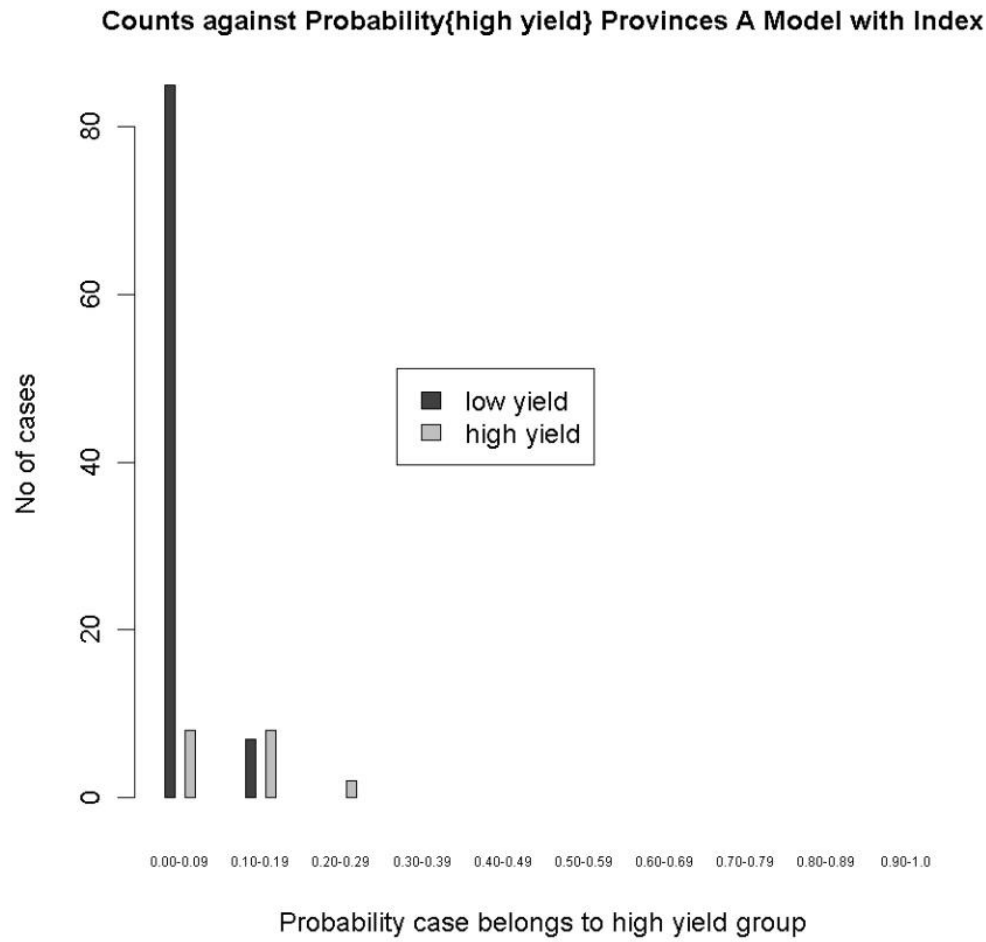


Figure 11.3 Probability bar chart for premises in new dataset when high yield is 5.5% or more

Since at the 5.5 per cent threshold 18 of the 110 cases were actually in the high yield group, the results produced were actually less accurate than those produced by the original model. They were also of no practical utility in investment decision-making.

Selecting properties for investment based on a ranking of the probability of high yield produced the results shown in Figure 11.4.

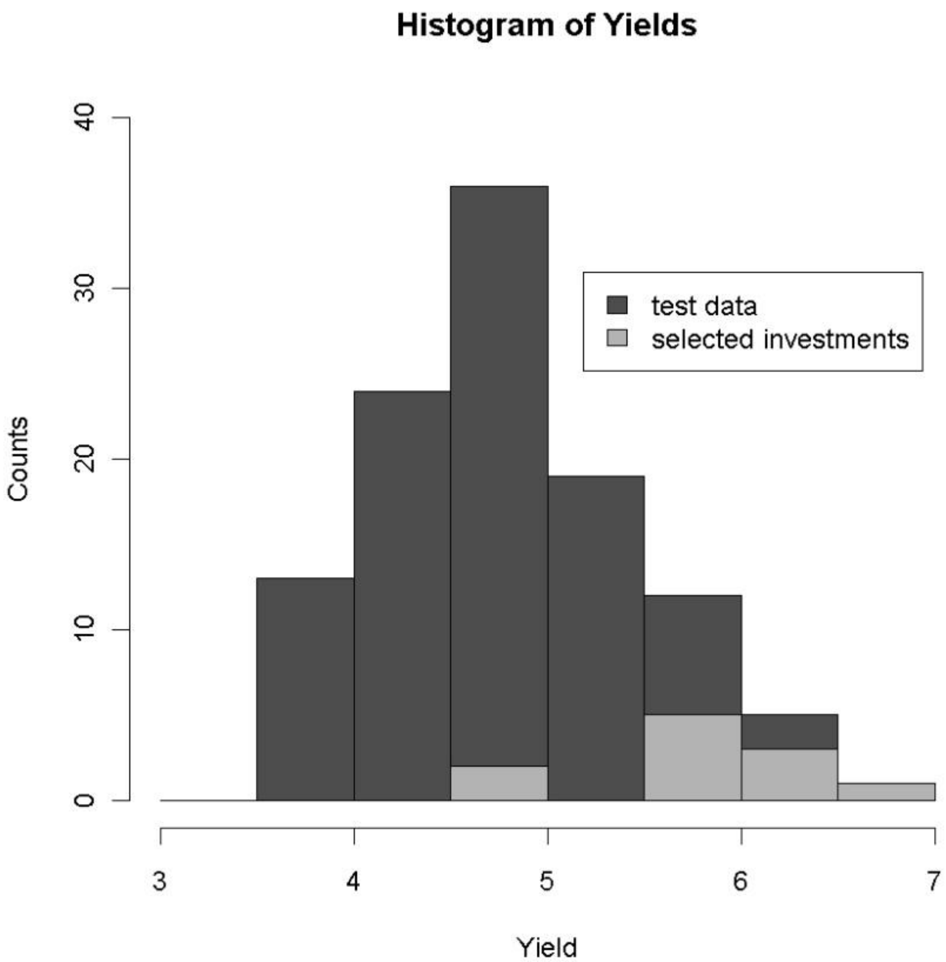


Figure 11.4 Histogram of yields for premises in new dataset when high yield is 5.5% or more

Again, the mean yield for the selected ten per cent of properties with the highest probability of being in the high yield group was 5.73 per cent. The mean yield for all the properties in the dataset was 4.82 per cent.

11.2.2 The ANOVA model

The problem with the logistic regression was that the yield was categorical: low or high yield groups and very little of the new data were in the high yield group. Therefore, insufficient lots could be identified for the purposes of building an investment portfolio based upon the high yield group.

Although the original analysis had been unable to identify a useful ANOVA model, the ANOVA did give an immediate prediction of the yield for each property. It assigned a continuous value for each lot. Ranking properties in terms of their predicted yield appears to be a reasonable approach to investment decision-making. Figure 11.5 shows a scatterplot of Predicted Yield against True Yield for the 110 properties in the new dataset based upon the indexed ANOVA model for *Provinces A* produced in section 10.3.3.1. The scatterplot shows that the relationships between predicted yields and actual yields are not very well defined.

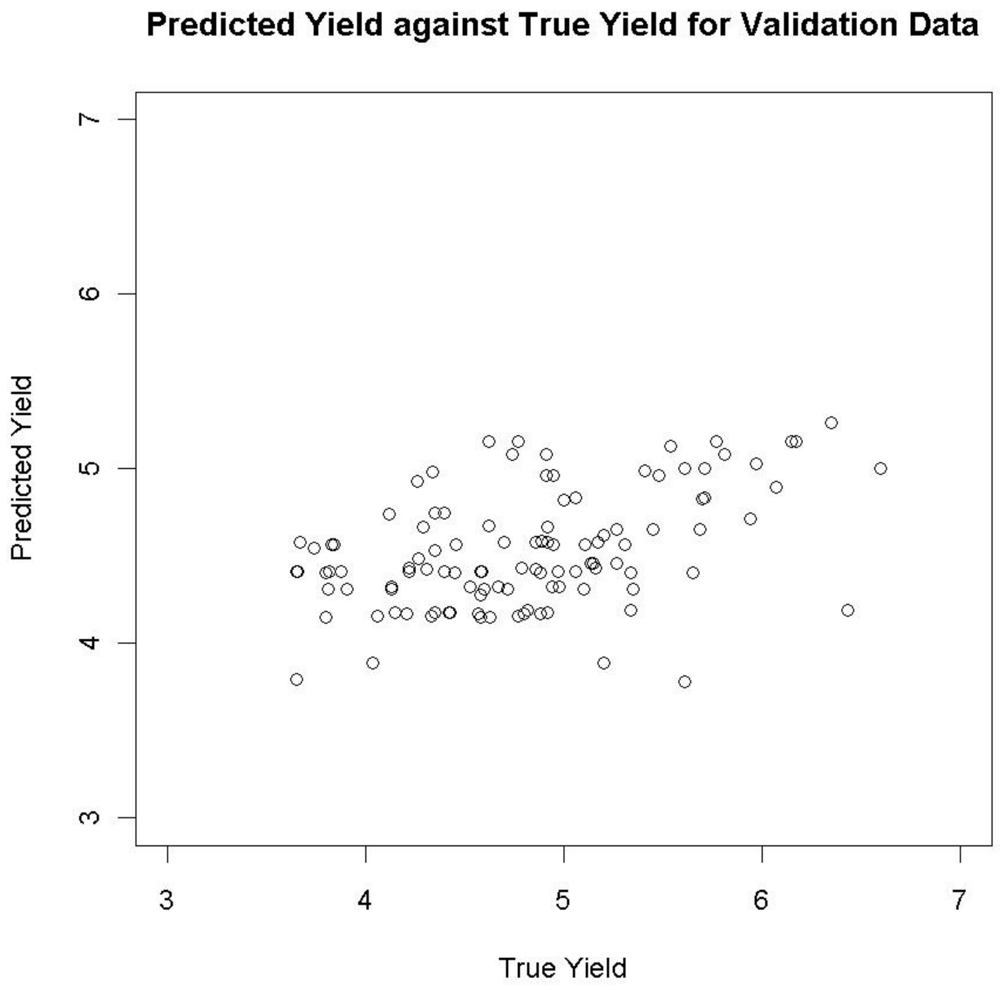


Figure 11.5 New data ANOVA predicted versus true yield scatterplot

Figure 11.6 shows the results of selecting the highest ten per cent of properties in terms of predicted yield. The mean yield for the eleven selected lots was 5.53 per cent compared to a mean yield of 4.82 per cent for the complete dataset. Selection according to this method therefore leads to a better investment portfolio than random selection of lots to purchase. The mean yield of the selected lots at 5.53 per cent was, however, lower than the mean yield for the lots selected using the logistic regression model, which was 5.73 per cent.

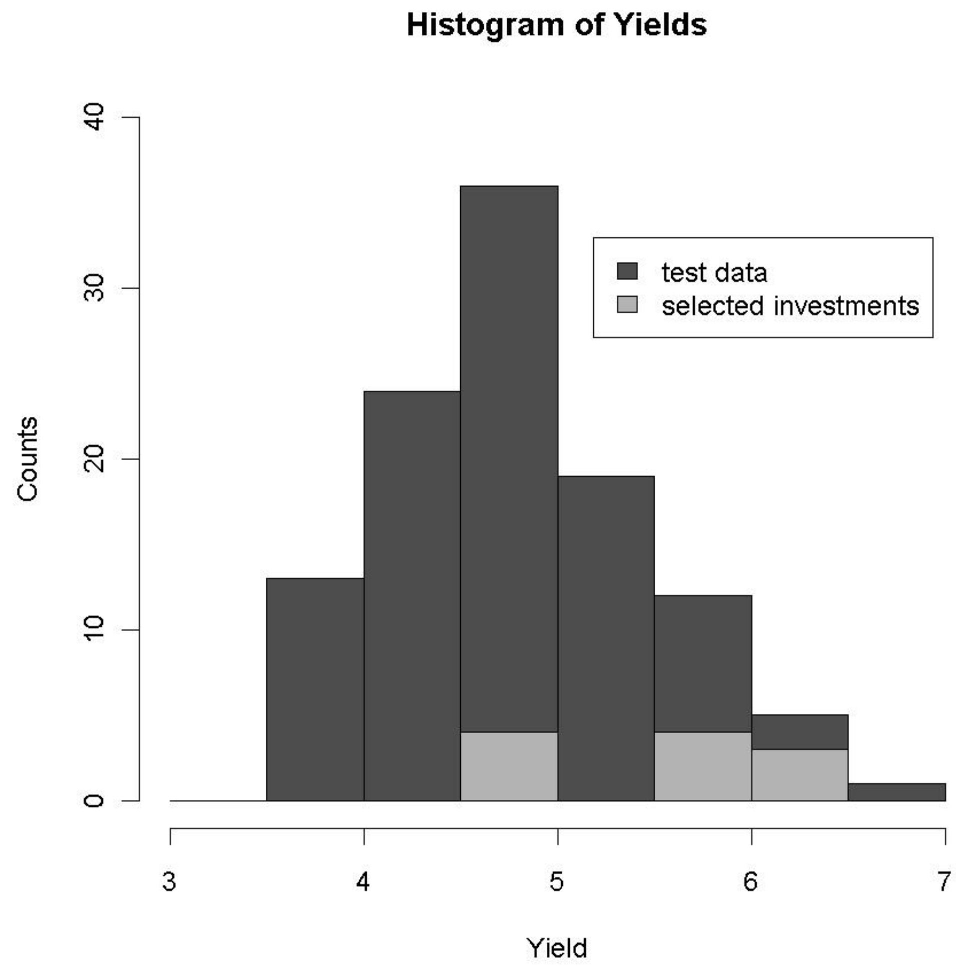


Figure 11.6 New data ANOVA histogram

11.3 Conclusion

The logistic regression indexed models, whether the threshold for high yield was set at 5.5 per cent or 6.35 per cent, produced a mean yield of 5.73 per cent for the ten per cent of properties selected as having the highest probability of falling in the high yield group. A lower threshold of 5.5 per cent for high yield group membership actually produced less accurate results. This contrasts with the ANOVA model, which produced a mean yield of 5.53 per cent for the ten per cent of properties selected as having the highest probability of falling in the high yield group.

The conclusion is that overall the logistic regression model is the preferred model. That is to say that the preferred model is the original indexed logistic regression model for *Provinces A* with selections based upon the ranked predicted probabilities of membership of the high yield group at 6.35 per cent threshold for the high yield group. The model already selected in the quantitative study is further validated by the post-study data for the period comprising October, 2006 until December, 2007 inclusive.

12 Conclusions and Recommendations

12.1 Introduction to the conclusions and recommendations chapter

Following the further validation using the post–study period data of the qualitative and the quantitative studies, the findings can be fully interpreted and discussed. Such interpretation and discussion can be undertaken on the basis that the research is very robust due to the triangulation of methodologies and a design that adequately addresses issues of validity and reliability.

The validated findings enable reflection in the context of a professional doctorate based upon extant literature. Hence, contributions to the researcher's work-place, to his own professional development and to the profession at large are capable of being considered.

12.2 Introduction to the summary of the findings

The interpretation is based upon the findings of the qualitative and quantitative studies. Taken together, the respective studies show that the initial yields of British banking-hall investment properties sold at auction are influenced by factors that are capable of being used in predictive models. Those factors identified as being most influential are capable of being incorporated into models which form the basis of a toolkit for the prediction of those retail bank premises likely to return a higher initial yield for portfolio building purposes.

The qualitative and quantitative studies found that the factors most likely to have influenced investors' decision-making and the yields of freehold retail bank premises were:

- Tenant bank company and brand
- Regional location
- Lot size as defined by hammer price
- Time

Taken together, the respective studies undertaken within the current research show that the following factors have the greatest influence on the initial yields of British banking-halls: tenant bank company; lot size; province as defined by super-region; and the macro-economic cycle whether represented by time or the Investment Property Databank (IPD) United Kingdom Retail Property Index. The research not only shows that these are the main factors affecting yield, but also provides the predictive framework that property investors and their professional advisors can use for assembling portfolios of retail bank premises that are likely to generate higher yields. This research was sponsored by companies employing the researcher. Markowitz (1983) stated:

> "The use of portfolio analysis by an individual investor can be desirable even if no one else in the world uses it."

Therefore, the analysis undertaken and the predictive framework built by the researcher are useful and a contribution to both knowledge and professional practice even if the researcher's sponsors are the only persons to rely upon them.

Models were produced both by logistic regression and ANOVA for time represented as either calendar time (*Time*) or the IPD United Kingdom Retail Property Index (*Index*). In all instances, logistic regression produced useful models for the data whether collapsed into the *Provinces A* or the *Provinces B* datasets. On the other hand, none of the ANOVA models proved to be of practical utility. This is a disappointment, because it was hoped that a useful ANOVA model might have been created with a view to predicting actual initial yield. Further validation from the post-study period supported the indexed logistic regression model for *Provinces A* as being the preferred one, as follows:

$$f(\lambda) \quad = \quad \lambda + \lambda_{Lk} + \lambda_{Pl} + \lambda_I I_i$$

All of the *Time* models produced by logistic regression included all of the linear terms that represented the linear factors tested. Some interactions between factors were also produced. In each form of collapse, the simplified models were preferred to the basic models. However, the *Index* model produced for both the *Provinces A* and *Provinces B* omitted the linear term *Bank*. This was not because *Bank* was not significant. Indeed, *Bank* was shown to be the next statistically significant term in the two *Index* models. A large population was used in the quantitative study. Statistically significant effects might not necessarily be practically significant (Agresti, 1996: 161-162).

It should be borne in mind that the qualitative study suggested that the linear term *Bank* was a significant factor influencing the initial yield. This qualitative finding is well supported by Enever and Isaac (2002). It was also supported by the part of the quantitative study using *Time* rather than *Index* as an independent variable. Furthermore, it needs to be appreciated that the quantitative study was founded on the data of premises transacted with respect to the three main banks involved in the sale-and-leaseback process. These three banks were also large in terms of their market share of British retail banking. Had the quantitative study been founded on data that included the smaller banks as well, the significance of *Bank* may well have been shown to have been greater. Moreover, the suggestion from the qualitative study is that *Bank* continues to be of practical significance to the professional practitioner.

The analyses suggested that the index model for *Provinces A* had greater predictive accuracy than the index model for *Provinces B*. Therefore, the logistic regression index model for Provinces A was the preferred of all the models. For this model, a predictive threshold at above 0.9 gave especially good results. However, the analyses suggested that at a threshold above 0.8, this model continued to provide a high positive prediction for the assembly of a banking-hall property investment portfolio. The 0.8 level was selected, because it gave an acceptable level of false positives and therefore provided a sufficient number of true positives. This model comprised the three linear terms:

- Lot size
- Index
- Provinces A

12.3 An overview of the significant findings

The main purpose of the current study was to identify ways in which investors purchasing freehold retail bank premises in Great Britain could maximize their rental yields. For the purposes of this study, initial yields have been adopted as the type of yield to be measured.

Hence, the core objective was to create a toolkit which comprised a predictive framework for identifying those banking-halls likely to produce the highest initial yields.

The specific research question was:

"How can property investors select freehold British banking-halls that are likely to provide the highest initial yield on their investment?"

Following on from this, it was hypothesized that the initial yield for banking-hall investments could be predicted statistically using a regression model, based on the following sub-hypotheses:

- There is some regional disparity in yields;
- Lot size does make a difference inasmuch as there is increased demand for smaller lots;
- Tenant banking company has an effect.

On the grounds that economic and business research considers the effects of fluctuations within the macro-economic cycle, the following sub-hypotheses were tested as being representative of that cycle:

- Time has an effect
- The IPD United Kingdom Retail Property Index

Others possible factors which were not tested by this research, but which potentially may warrant investigation are (see also sections 12.7 and 12.8 for the reasons why these factors were not tested):

- The less reversionary banking-halls are in especially high demand from investors;

- Premises let to the smaller banks and demutualized building societies do not necessarily command higher yields that those let to the main retail banks.

The variables where the sub-hypotheses were found to exist were those tested in the quantitative study. The factors found to have the greatest influence on the dependent variable, initial yields of British banking-halls sold at auction as property investments, were the following independent variables:

- Tenant banking company;
- Lot size in terms of price paid;
- Province as defined by regions being grouped together into super-regions;
- The macro-economic cycle.

These findings are tempered by the variable *Bank* having been omitted purely on the grounds of practical significance from the preferred model (the logistic regression *Provinces A Index* model).

Nevertheless, for these independent variables, the sub-hypotheses, and hence the main hypothesis, are proven.

With respect to lot size, the question was posed at the outset of the study whether smaller lot sizes were in more demand due to competition from smaller, private investors. The study found that larger lot sizes generally attracted higher yields. This supports the hypothesis that investors will pay more *pro rata* for smaller lots, which attract a higher years' purchase and therefore a lower All Risks Yield than larger lots would.

These findings are capable of forming the basis of a predictive framework for identifying those banking-halls likely to produce the highest initial yields. The Conceptual Framework for Predicting Yield, shown on page 84, has been confirmed. Citing Miles and Huberman (1984: 3), Leshem and Trafford (2007: 289) show how such a framework may evolve during a research project. During the current study, the framework has been modified by the removal of those factors not tested and by the change of *Region* into *Super-region*. For the sake of clarity, this study calls super-region *Province*. The modified framework is entitled, *Validated Framework for Predicting Yield*, and is shown in Figure 12.1.

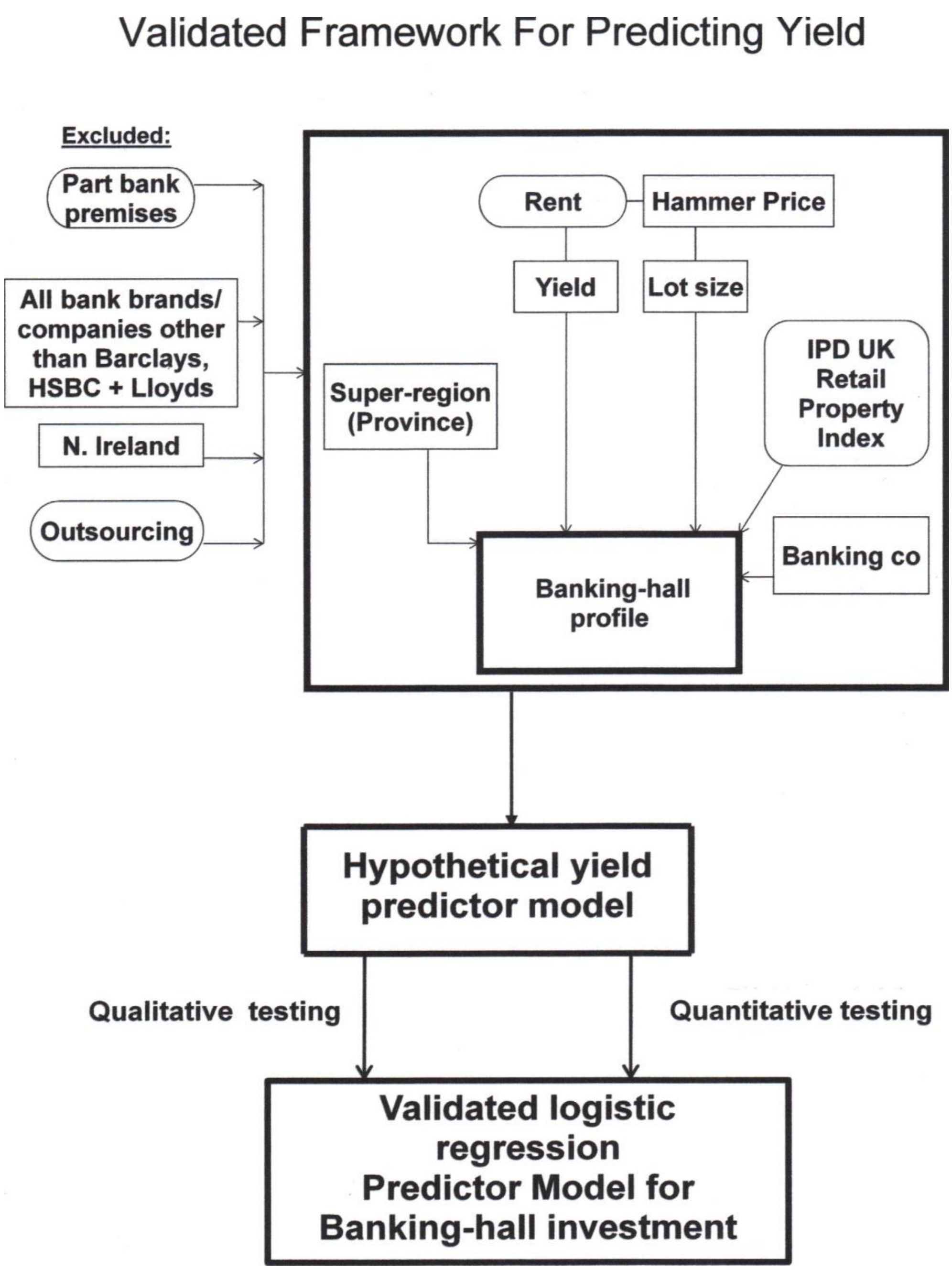


Figure 12.1 Validation of the framework for Predicting Yield

12.4 Consideration of the findings in the light of existing research studies

Extensive searches on the electronic databases failed to find any other studies into British banking-halls as a property investment either for the current study period or an alternative one. However, it was possible to draw on research into factors affecting the yields of other types of property. Some of these studies were confined to the United Kingdom. Others were on property abroad. Since retail bank premises are normally treated as a sub-type of retail property (Jones Lang LaSalle, 2008b and 2009b), the current research focused on the findings of other studies into retail property.

Ambrose and Nourse's study established the need to recognize the distinct types of property rather than to treat real estate as a single asset class (Ambrose and Nourse, 1993). Since 1993, retail property has been treated as a type of property (Ambrose and Nourse, 1993; McGregor and Schwann, 2003; Chen *et al.*, 2004). Banking-halls are a distinct sub-type of retail property. This is recognized by the auctioneers listing banking-halls as such a sub-type.

Previous American studies have identified region as being an influence on yield (Ambrose and Nourse, 1993; Nelson and Nelson, 2003; Cheng and Roulac, 2007). However, such a regional effect on retail yields in the United Kingdom is not supported when region is applied as defined by Government Office Regions. However, some (Hoesli *et al.*, 1997; Guy and Hennesberry, 2000), but only some, sources support the hypothesis that there is a super-regional effect on yield. These were based upon a north-south-divide. The analyses in the current study suggested a super-regional, or *Province*, effect based upon a north-south divide, but with London and the south-east of England achieving the lowest yields. In contrast, the findings of the current study suggest that the super-region furthest from London, however created from combining regions, achieves the highest yields.

Ambrose and Nourse suggest that in the United States lot size does have an influence on yield. Baum *et al.* (2006: 76) and Whitmore (2007) suggest that in the United Kingdom yield is influenced by lot size. Those findings are borne out with respect to the lot sizes of British banking-halls when lot size is measured by the Stamp Duty Land Tax thresholds.

Although no studies into the effect of specific tenant company on yields of retail bank premises were identified, Enever and Isaac (2002) cite specific companies within a business sector as having different share pricings. They suggest that differences should be considered as a reflection of specific risk when share investment portfolios are assembled. If this reasoning is extended to different tenant companies in the retail banking sector, then this is supported by the findings of the current study. The arguments of Enever and Isaac suggest that, in accordance with the qualitative study, *Bank* should have an influence on initial yield. In parts of the quantitative study, *Bank* was shown to be both statistically and practically significant. It was shown not to be practically significant in the preferred model. However, due to the nature of the available data, the quantitative study relied upon the data of the three main banks. Had it been possible to have conducted the quantitative study using data from both the three main banks and other banks, then the variable may have been shown to be practically significant in the selected model.

The findings of the studies of Krystalogianni and Tsolacos (2004) and Dunse *et al.* (2007) that property investment yields are affected over time are borne out. This is a reflection of the macro-economic cycle and market risk fluctuating over time (Tipping and Lam, 2010). In the current study, time, whether represented by calendar time or the IPD United Kingdom Retail Property Index, was shown to influence the yields of British banking-halls.

Previous studies, considered in Chapter 4 (Influencing factors of banking-hall investment: theoretical perspectives), support and underpin the findings of the qualitative and quantitative studies that form part of the current research.

12.5 Implications of the study for current theory

Table 12.1 shows the number of British retail bank branches by the six main banking groups. Between them, these groups still retain 10,247 branches. This is very similar to the number of branches existing in 2004 as shown in Table 1.1. Clearly, the retail branch is still alive. Investors in these branches still need a predictive framework.

Table 12.1: Main branded bank branch networks in Great Britain in 2011

Bank	**Number of branches**
RBS/NatWest • Royal Bank of Scotland (644) • National Westminster (1,545)	2,189
Barclays	1,658
Lloyds Banking Group • Lloyds TSB (1,785) • Halifax (669) • Cheltenham & Glos (164) • Bank of Scotland (295)	2,900
HSBC	1,300
Nationwide • Nationwide B.S. (700) • Other brands (100)	800
Santander	1,400
Total	**10,247**

Source: Kivlehan, 2011.

The findings of this study extend extant theory. Extant theory relates to retail property as a main property type. Theory can now be extended to retail bank premises in Great Britain. The theory that now exists with respect to the main factors influencing the yields of British banking-halls can now be used as part of the practitioner's toolkit. The quantitative analyses have shown that the logistic regression models are useful ones that are capable of being used in an empirical way for predicting lots which are likely to fall in the higher yield group. However, such an empirical approach necessitates running the data through a computer using the appropriate script. Such a method is time consuming and is not always practical. This is especially so for those lots that are sold at auction. Most British banking-halls sold as property investments are sold by auction. The generation of theory empowers the practitioner with insight and tacit knowledge that can be used to make informed decisions (Drake and Heath, 2011). Having extended extant theory on retail property to banking-halls, it is now possible to apply such theory to the real world. Thus, investors and their professional advisors can utilize the theory in the context of the auction arena. The theory derived from the study is part of the toolkit available to investors and their professional advisors.

Having a toolkit for investors and their professional advisors remains useful, because retail bank premises retain their appeal to investors (Duxbury, 2010). Duxbury argues that the experience of the credit crunch has shown that the main banks have been too important for the government to permit their failure. As a result, British banking-hall investments continue to provide a popular and secure form of investment to private investors (Duxbury).

12.6 Findings that appear to fail to support or only partially support the hypotheses

The preferred model, on the grounds of having the greatest predictive power, omitted banking company as being one of the main influencing factors on yield. However, this was not because banking company was not significant. It was shown to be the next most significant factor in statistical terms. It was merely omitted from the model on the grounds of practical significance. Had it been possible to analyze data with respect to the other banks as well as the main three, then banking company might have been shown to be of practical significance in the preferred model. It could also be argued from the findings of the qualitative study, professional practitioners consider this variable to be of practical significance.

One hundred per cent of the respondents in the qualitative study stated that region was one of the factors most influencing the yield of British banking-halls. The respondents were all professional staff working in the main auction houses selling banking-hall investments. Those auction houses base their definition of region on that adopted by the Government Office Regions adopted by central government. However, such an influence on yield based upon region defined in this way is in contrast to findings of both the theoretical perspectives and the quantitative study that each comprise part of this research. The quantitative study confirms the findings of the theoretical perspectives that there is a regional effect, but on the basis of super-regions (called *Provinces* in this study). This study shows that one of the main influences on yield is a super-regional one based upon a north-south divide.

12.7 Limitations of the study

The research was confined to an investigation into freehold banking-halls. This was for two reasons. Firstly, the sponsors of this study clearly expressed that they were only interested in research into freehold retail bank premises. Secondly, the dearth of leasehold premises sold during the study period meant that there were insufficient leasehold cases to populate enough cells in the contingency table to permit meaningful analysis. Therefore, the findings of this study should not be generalized beyond freehold properties or in Scotland beyond feuhold premises.

Collapsing the number of cells from five thousand seven hundred and sixty to one hundred and forty-four rather than to ninety-six in the contingency table may still have permitted meaningful analyses. However, the generation of the histogram output of yield with respect to the data showed very graphically that the distribution of yield was bimodal. Therefore, to facilitate meaningful analysis of the dependent variable, yield was collapsed into two rather than three categories. This meant that the models built were limited to predicting whether a particular lot would be likely to fall into the high or low yield category.

The failure to develop a useful ANOVA model meant that it has not been possible to extend the toolkit further to provide a means of assessing the actual yields, which could then be used to capitalize rents to provide a bid price.

Due to the need to reduce the number of cells in the contingency tables to facilitate meaningful analyses, those freehold properties analyzed were restricted to those let to the three large banks leasing the largest number of premises sold during the period. Such a restriction helps to considerably reduce the number of cells in the contingency table whilst retaining seven-ninths of the premises. The analyses are based on premises let to Barclays, HSBC and Lloyds TSB. The findings arising from the quantitative study are capable of generalization to other retail bank premises let to these three banks. However, questions of validity might arise if an attempt were made to generalize the findings to premises let to other banks. The three leading experts, questioned during the semi-structured interviews, stated that investors preferred to purchase premises let to the main banking brands rather than those let to secondary retail banks. This suggests that the findings of this study should not be generalized to premises let to the other banks.

At the outset of the study, it was suggested that the hypothesis that the less reversionary banking-halls are in especially high demand from investors should be tested. The three leading experts had suggested during interviews that investors sought the less reversionary banking-hall investments. This was supported by half of the respondents to the questionnaire suggesting that the unexpired lease length might have an influence. However, these respondents did not specify the optimal length, but it was considered that longer leases were preferable to shorter ones to investors. The lots subjected to the quantitative study did not all have the same length of unexpired lease. This might have an effect on yields even if not significant in practical terms. Nevertheless, the potential influence of unexpired lease length should be borne in mind.

It was not possible to test whether premises let to the smaller banks and demutualized building societies attracted the same or higher yields as those let to the main retail banks. Again, this was due to issues relating to the scarcity of data. Especially in the light of the semi-structured interviews suggesting higher demand for premises let to the main banks, the findings of the current study should not be generalized to premises let to the smaller banks and demutualized building societies.

Although it has not been subject to scrutiny in this investigation, there is some anecdotal evidence that there may be for any given vicinity a minimum rental value applicable to a banking-hall irrespective of its size. This arises out of the need for banks to have a physical presence in a locality. Since the supply of retail banking premises tends to be fixed in the medium term, demand is likely to keep rents at or above a minimum level irrespective of floor area. This can be likened to the demand for kiosks from certain classes of traders in certain circumstances and at certain locations. This is highlighted by Reuveni (2002) in his research into newspaper vending in Germany prior to 1933.

Reuveni traces the process of the development of newspaper kiosks in Germany from the 1870s, when the firm of Brockhaus started selling reading material at railway stations in Leipzig. Georg Stilke became the best known firm to follow this model. In 1882, it acquired the rights to sell reading material at railway stations in Berlin. Its business soon expanded to railway stations throughout north Germany. Also, the firm was able to expand its kiosks throughout the Berlin underground railway system when that service commenced in 1902. According to Reuveni, research indicates that Georg Stilke came to control forty per cent of such kiosks throughout Germany. By the end of the nineteenth century, some two hundred operators sold reading material at railway station kiosks, and following the First World War the number stood at two hundred and fifty operators. This business was conducted at more than eight hundred railway stations (Reuveni, 2002). Hence, the competition from both the large firms and the independent operators was keen. Such competition inevitably kept rents high so that news-vendors could retain a presence at these locations. Indeed, Reuveni states that a quick turnover was required due to the high rents paid for kiosks.

12.8 Recommendations for further research

Hypotheses suggested at the outset of this study, but not tested within it, should be tested to broaden knowledge, subject to there being sufficient data to enable this. These are:

- The less reversionary banking-halls are in especially high demand from investors;

- Premises let to the smaller banking companies and demutualized building societies do not necessarily command higher yields that those let to the main retail banks.

The first of these two hypotheses was not tested due to insufficient data. A minority, albeit a sizeable minority, had cited lease length as being a relevant factor. However, these respondents gave no clear indication as regards to what length of lease. Aside from the questinon of having had to collapse the data, the second of these hypotheses was not tested due to there being no data available.

The results of the semi-structured interviews within the current study, in particular, suggest that premises let to the smaller banking companies and the demutualized building societies may well command higher yields than those let to the main retail banks.

Also, those factors explicitly placed outside the conceptual framework, either by design or due to insufficient data should be investigated, subject to there being the available data. These factors are:

- Properties only partly occupied as bank premises

- Other bank brands and companies

- Banking-halls in Northern Ireland

The influence of unexpired lease length was not investigated, because only 16.7 per cent of the respondents in the qualitative study supported it.

Following Reuveni's (2002) research into rental demand for early twentieth century German news-vending kiosks, further research is recommended to investigate whether there is indeed a minimum rental value attributable to banking-halls for a given locality irrespective of floor area. Additional research should conclude whether or not banks will pay over and above the going rate per square metre in a locality merely to secure one of a very limited number of banking-halls in that locality. Can the demand to rent banking-halls in the present-day United Kingdom be likened, for example, to the demand to rent news-vending kiosks at German railway stations between 1870 and 1933?

Leasehold investment interests in banking-halls were excluded from this study specifically because the sponsors had requested a study into only freehold premises. Notwithstanding that, leasehold premises would not have been capable of being tested in the quantitative cross-sectional study used in the current research due to there having been far too few leasehold cases to sufficiently populate the cells in cross-tabulations. Leasehold interests should be investigated, subject to there being sufficient data to facilitate a useful study.

The predictive model is based upon the data from those three main banks that sold the freehold interests of the largest number of British banking-halls. This limits the ability to generalize from the findings and the model. It is, therefore, suggested that research is undertaken on the other and smaller banking companies in order to make the model more robust and applicable to those other banks.

The current study is related to retail bank premises, which have been identified as a sub-class of retail property. Further research may be undertaken to develop predictive frameworks forecasting yields for investment in other retail sub-classes.

12.9 Identifying other knowledge gaps

At an early stage of the current research, the candidate sought to adequately define the dependent variable *yield*. This was defined as being initial yield in accordance with practice in the candidate's workplace, the auction market-place and the auctioneers' catalogues and promotional literature. In defining *yield* as initial yield, the candidate discovered that the topic merited a conference paper in its own right. In researching the paper, the candidate was able to ascertain that there was a great deal of lost history with respect to the development of the investment valuation tables and that there was a lack of adoption of effective yields.

The investment valuation tables can be shown to have evolved over millennia (Tipping, 2006). However, much of the knowledge appertaining to the long history is in danger of being lost permanently due to the data being recorded in old documents that either have been lost and forgotten or are very fragile and in danger of disintegration. This is an area where the combined interests of history and valuation could be applied to recover and record that knowledge for the future.

The property investment auction market uses initial yields. Initial yields for freehold and other investments in perpetuity are simple to calculate. They, therefore, lend themselves for use in a fast moving auction. Such initial yields are calculated on the basis of nominal yields (Allsop & Co., 2004). Nominal yields are calculated on the assumption that rent is payable yearly in arrears. However, the reality is that rents are paid in advance. The rents for the banking-halls in the current study are paid quarterly in advance (propertyauctions.com, 2004). Some other rents are paid monthly in advance. Especially in the aftermath of the credit crunch, there has been growing pressure from some tenants for rents to be paid monthly in advance (Blackhurst, 2008). Rental payment patterns impact on capital values due to the time value of money (Tipping, 2006). More advanced valuation methods account for such impact through the use of effective yields. Arguably, the lack of uptake of effective yields is more due to a gap in professional practice rather than a gap in knowledge. However, there is a case for making the use of effective yields more attractive for practitioners to use them more. This is a gap in the knowledge about how the use of effective yields could be applied in a very fast moving auction. It remains to be seen if that gap can be filled in a practical way.

12.10 Summarizing the Professional Doctorate

Following the interpretation and discussion of the findings of the research, it is appropriate to include a section examining the impact on professional practice and any observations and recommendations (Murray, 2006: 199 and 226). The Professional Doctorate should display some attributes that have gone beyond those of some other forms of doctorate. In particular, the Anglia Ruskin University Research Degrees Regulations (Anglia Ruskin University, 2010: 67) state that the Professional Doctorate needs to:

> "Display appropriate evidence of originality and independent critical judgment AND constitute a contribution to professional practice AND exhibit development of professional competencies."

Since the current study displays evidence of originality and independent critical judgment, the following sections of the current chapter focuses on the contribution to professional practice and the development of professional competencies. In short, the Professional Doctorate should show an impact on professional practice. In a more generic sense, such impact can be summarized as follows:

- Contribution to professional activity of the researcher in the workplace.
- Contribution to the researcher's own professional development.
- Contribution to the profession.

In terms of the current study to build a predictive framework for the investment yields of retail bank premises, the anticipated impact on professional practice has been summarized in Chapter 1 as follows:

- Contribution to the professional advice activity for banking-hall investment clients in the researcher's company.

- Contribution to the researcher's own professional development in property investment knowledge.

- Contribution to the surveying profession by providing new, special skills of banking-hall investment for the retail property market. Some research output has already been disseminated by the candidate in the form of published works listed in Appendix XI. Specific findings from the current research will be disseminated in either conference or journal papers. There is also scope for collaboration with the Royal Institution of Chartered Surveyors for the provision of guidance notes.

12.11 The role of the Professional Doctorate

Murray (2006: 38-40) along with Heath and Drake (2011: 1-4) review extant literature about the distinctions between doctorates of philosophy and professional doctorates. Each show that convention has it that the distinctions between the two types of doctorate are founded on academic and professional research respectively. However, they also show that the perceived distinctions have in some instances become blurred.

In seeking to define the distinction between doctorates of philosophy and professional doctorates, Murray cites Doncaster and Thorne (2000: 392) who defined the distinction on the basis of *professional scholar* and *scholarly professional*. Hence, the doctorate of philosophy is seen as an introduction to an academic career, whereas the professional doctorate is perceived as a means of improving professional practice (Murray, 2006:38). Thus, in addition to creating new knowledge, the professional doctorate can be expected to have an impact upon the practitioner's workplace and contribute to new knowledge within the profession (Murray, 2006: 38; Drake and Heath, 2011: 90).

Notwithstanding the distinctions made by Doncaster and Thorne about the two types of doctorate, Drake and Heath draw upon a number of recent studies to argue that the real distinctions are in fact very little more than cosmetic. The thrust of their argument is that apart from practitioners undertaking professional doctorates continuing to remain in practice during the programme, the outcomes of the two doctorates are not too dissimilar. They suggest that although practitioners changed the conduct of their practices as a result of practitioner research, the impact within the profession as a whole appeared scant.

Notwithstanding the caveats made by Drake and Heath, the current study undertaken within the context of a professional doctorate programme not only makes an original and independent contribution to knowledge, but it has an impact on profession practice. Such impact on professional practice takes the form of changes in the researcher's practice and the dissemination of knowledge throughout the profession through publication.

12.12 Contribution to the professional activity of the researcher in the workplace

In a study conducted by them in 2008, Drake and Heath (2011: 90) found that the greatest impact on professional practice was felt in researchers' own workplaces and professional practices. Moreover, they found that the greatest impact was in the way that practitioners reflected and approached their own profession practices.

The present study and, more especially, peer-reviewed papers presented and published by the candidate during the professional doctoral programme have resulted in changes to the candidate's professional practice. As a result, the scholarly professional is now more reflective and approaches the conduct of practice at a higher level (Murray, 2006: 38; Drake and Heath, 2011: 90-92 and 96). The changes to the candidate's professional practice can be summarized as follows:

- Professional practice is approached in a reflective way

- Problem solving is considered in a similar way to researching and writing papers
- The candidate's own peer-reviewed papers have been drawn upon as a way improving practices in his own workplace
- The findings of current study have been used in selecting banks to bid for
- Literature reviews have been used to look at other property types, but qualitative studies have not been undertaken

12.13 Contribution to the researcher's own professional development

The current candidate observed that to some extent, the reservations that Heath and Drake have about personal professional development on professional doctoral programmes were borne out on his own programme. In addition to the arguments of Drake and Heath, the candidate observed a great reluctance amongst the majority of fellow students to publish their own papers. Murray (2006: 39) suggests that professional doctorate programmes offer opportunities for publishing papers and that some programmes have specific modules for doing so. The current candidate has observed that to date there has been a correlation between the publication of papers and success on the programme. On the basis of this, he has long stated the case for making the publication of papers in journals and at research conferences a mandatory part of the programme.

The current candidate has published several papers during his time on the professional doctorate programme. The original purpose of so doing was to be able to cite his own peer-reviewed work in this thesis. However, he has since come to view such publication as an important part of his own professional development. Moreover, since then he has started to publish outside the area of the current research as a way of extending his professional development further. First and foremost, the researching, writing and the presentation of the papers ensure that the candidate fully reflects upon the issues and develops an understanding of the topic at a much higher level. Also, the feedback from peers is constructive.

The four papers researched and published by the candidate during the professional doctorate programme are shown in Table XI.1 in Appendix XI. The three papers related to the current study were all peer-reviewed. The paper entitled *Sale-and-leaseback as a British Real*

Estate Model was published twice in 2007. Having been peer-reviewed it was first presented at the FIG (Fédération Internationale des Géomètres) 2007 Congress in Hong Kong. Shortly after that, the editor of the *Journal of Corporate Real Estate* asked if she could publish it again in that journal. After the paper was double-blind peer reviewed, the candidate added a further paragraph of his own volition after reflection before it was published in that academic journal.

Following the publication of the papers listed in Table XI.1 in Appendix XI, the candidate has a better understanding at a higher level of thinking about the issues being covered. He has learnt how to be more critical. In addition to the peer-reviewed papers relating to both the candidate's professional practice, the paper entitled *Identifying Clay-construction Buildings in a Norfolk Market Town* was researched and presented primarily with a view to demonstrating, both to the candidate and others, that he was capable of participating in research projects beyond his own field. The candidate has acquired the skills and confidence to write, present and publish papers at conferences and in journals. Since the publication of papers enables the candidate to further develop his own professional knowledge, he will continue to publish papers.

12.14 Contribution to the surveying profession

Contribution to the profession can be measured in different ways and at different levels. The purpose of the current research is based upon the sponsors' need for a toolkit comprising a predictive framework for placing those banking-halls expected to achieve the highest yields into an investment portfolio. Again citing Markowitz (1983):

> 'The use of portfolio analysis by an individual investor can be desirable even if no one else in the world uses it.'

It can be argued that just as it is acceptable for the findings of this research to be used by only one investor, it is similarly acceptable for any contribution to the surveying practice to be of use to only a small number of practitioners. Hence, any contributions to the profession arising from either the findings of the current study or other papers written by the candidate are worthy no matter how few practitioners rely on those contributions.

All of the papers published by the candidate during the current research have been placed in the public domain for use by others. It has been ascertained with respect to all of them, save the most recently presented one, that they have been used and cited by others either in professional practice or research.

The first paper, *The Impact of the Time Value of Money on Valuation Practice*, is cited in the footnote of a paper written in Russian, the title of which translated into English is *History of Methodology in the Professional Valuation of Capital*. The paper was published in November, 2009. The author of that paper is Andrey Igorevich Artemenkov, who is of the State University of Management and also of the Russian Society of Appraisers.

The first paper is also cited by Fornero (2007) in *Cronología Fotográfica De Las Finanzas*. Although the latter publication is written in Spanish and published in Argentina in 2007, the element relating to the investment valuation tables during the seventeenth, eighteenth and nineteenth centuries clearly draws heavily on the current candidate's work. The candidate researched the history of the investment valuation tables at reference libraries in the United Kingdom using a hermeneutic methodology.

The second paper, *Sale-and-leaseback as a British Real Estate Model*, was initially presented at the FIG 2007 Congress in Hong Kong in May, 2007. It had been peer-reviewed for the FIG. At the invitation of the editor of the *Journal of Corporate Real Estate* five weeks after the FIG congress, the paper was submitted for double-blind peer review with a view to having it published in the journal. Some more academic citations were added to satisfy the first reviewer. The second reviewer wrote that it was the best paper that he or she had ever reviewed for a journal and that it had enabled him or her to resolve a problem in the Netherlands. It has been ascertained that the paper has been cited in at least seven other academic papers including ones published in Sweden (Muyingo, 2009; Björklund, 2011), Malaysia and Australia. It has been cited in peer-reviewed work in the Journal of Corporate Real Estate (Omar, 2010).

The second paper has also been used as a teaching aid at another university. The current candidate received e-mails from two Master of Science students at Oxford Brookes University stating that the paper was being used as a teaching aid on their real estate Master of Science course. Whilst at the FIG 2010 Congress in Sydney, the current candidate discussed the matter with an Oxford Brookes lecturer, who had once been his tutor as an undergraduate at the former Portsmouth Polytechnic. The Oxford Brookes lecturer confirmed that he was familiar with the paper and knew that it had been used as a teaching aid at the university by one of his colleagues. Furthermore, a Ph.D. student researching finance at City University, London contacted the current candidate for permission to use the paper in his own research.

The third paper, *Identifying Clay-construction Buildings in a Norfolk Market Town*, is outside the sphere of the current study and was not peer-reviewed. However, it was written and presented, because the candidate wished to demonstrate that he could use different research methodologies and techniques, and because it was in a field in which he wishes to advance his knowledge. The paper was one of two featured in an article by the chairman of the Property and Facilities Management Division of the Hong Institute of Surveyors in *Surveyors Times* (Yeung, 2010). Also, the current candidate was approached in early 2011 by an undergraduate of Anglia Ruskin University Department of the Built Environment, who had been given the paper as starting point for his own dissertation on new clay-construction buildings.

Drake and Heath (2011: 99) state that the validity of a piece of research can be measured by the extent of its peer acceptance. It therefore follows that if that research can be underpinned by the researcher's own peer reviewed papers, then the research is more valid. The candidate's contribution to the surveying profession through publication has, therefore, given the research a validity that extends beyond the validity associated with empirical data.

12.15 Final conclusions

The research design ensures that the findings of this research remain very robust through a design to limit threats to validity and reliability and through the triangulation of methodologies. As a result, findings are made which make an original and independent contribution to knowledge. A gap that had been identified in the knowledge is filled. This is highlighted. This chapter gives an overview of the significant findings of this research before considering them in the context of extant studies and theory.

Not only has the candidate's research shown originality and independent critical judgment, but it also makes a contribution to professional practice and a contribution to his own professional development. It has also made an original and independent contribution to knowledge. Furthermore, it has made a contribution to the candidate's own professional practice. A contribution has been made to professional activity within the researcher's workplace, to the researcher's own professional development and to the profession at large through the presentation and publication of papers at conferences and in an academic journal. Further contribution will be made through subsequent publication. In the meantime, the contribution to the profession through publication has further improved the validity of the current study.

This chapter also identifies other gaps in knowledge and practice, especially those in how practitioners treat and define yield. Those are gaps which can be later investigated by the current researcher and others. However, these gaps are more gaps in professional practice rather than gaps in knowledge.

Limitations are highlighted in this final chapter. In particular, the collapse of data needed to render them capable of meaningful analysis meant that the findings of the statistical analyses were confined to the three banks with the most premises sold at auction.

The selection of the indexed logistic regression model for *Provinces A* as the preferred one was supported by additional validation from the post-study period.

Recommendations for further research are made. These include hypotheses not tested in the current study and other areas explicitly placed outside the conceptual framework. Findings from these areas might be capable of extending the validated predictive model within the predictive framework even further. In which case, it would make the model more robust and more useful.

However, this study fills a gap in the knowledge and provides a toolkit comprising a predictive framework that property investors and their professional advisors can use for assembling portfolios of those retail bank premises that are likely to generate higher yields.

References

€100m Sony leaseback (2005) *Estates Gazette*, 15 October, 541, p. 39, London: Reed Business Information.

£46.6m taken at C&W' (2006) *Estates Gazette*, 9 December, p. 69, London: Reed Business Information.

A Society of Gentlemen in Scotland, (1771), 'Annuities', *Encyclopaedia Britannica,* pp. 318-323, Edinburgh: Colin MacFarquhar.

Abbey falls into £2m VAT trap (2006) 'Abbey falls into £2m VAT trap', *Estates Gazette*, 8 July, p.36, London: Reed Business Information.

Ackerman takes Liberty 2 from RBS (2006) 'Ackerman takes Liberty 2 from RBS', *Estates Gazette,* 16 December, p.14, London: Reed Business Information.

Adams, A., Booth, P., Bowies, D. and Freeth, D. (2003), *Investment Mathematics,* Chichester: Wiley.

Adams, A.T. and Clarke, R. (1996) 'Stock market reaction to sale and leaseback announcements in the UK', *Journal of Property Research,* 13, pp.31-46.

Adams, C . (2004) 'Sale/Leaseback Can Unlock the Value of Your Real Estate', *Expansion Management*, October, 19(10), p.4, Business Premier Source.

Afraid to ask? Outsourcing (2005) 'Afraid to ask? Outsourcing', *Estates Gazette*, 29 January, p.80, London: Reed Business Information.

Agresti, A. (1996) '*An Introduction to Categorical Data Analysis'*, New York: John Wiley & Sons, Inc.

AIB confirms sale of 12 Dublin branches (2006) *Business World*, September. Available at: http://www.businessworld.ie/cgi-bin/printer_friendly?a=1499282 [Accessed: 1 November, 2008].

Akaike, H. (1974) 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control,* 19(6), pp.716-721.

Allan, J., Murphy, J. and Kelly, T. (1996) 'Emissions in the heavy clay industry', *Key Engineering Materials,* 118, pp.249-256.

Allied Irish Banks plc (2005), *Be at AIB at the Post Office*, Republic of Ireland: Allied Irish Banks.

Allsop unloads sale and leasebacks (2006) 'Allsop unloads sale and leasebacks', *Estates Gazette*, 15 July, p.65, London: Reed Business Information.

Allsop & Co. (2004) *Allsop & Co. Website*, June, 2004. Available at: http://www.allsop.co.uk

Allsop & Co. (2009) *Commercial auction results.* Available at: http://www.allsop.co.uk

Amaratunga, D., Baldry, D., Sarshar, M. and Newton, R. (2002) 'Quantitative and qualitative research in the built environment: application of "mixed" research approach', *Work Study,* 51(1), pp.17-31.

Ambrose, B.W. and Nourse, H.O. (1993) ‘Factors influencing capitalization rates’, *Journal of Real Estate Research*, 8(2), pp.221-238, American Real Estate Society.

Anglia Ruskin University (2010) *Research Degrees Regulations*, Chelmsford: Anglia Ruskin University.

Archambault, S. (2000) *One-Way ANOVA*, Wellesley. Available at: http://www.wellesley.edu/Psych205/anova.html [Accessed: 16 May 2008].

Army housing brings in £4bn (2004) *Estates Gazette*, 24 July, p.31, London: Reed Business Information.

Asson, T. (2002) ‘Real estate partnerships: A new approach to corporate real estate outsourcing’, *Journal of Corporate Real Estate,* 4(4), pp.327-7.

At your service: the Post Office (1995) *The Economist (US)*, 28 October, p.66.

Badley, G. (2010) *Write a lot better*, INSPIRE Anglia Ruskin University website, pp.1-27, Cambridge: Anglia Ruskin University. Available at: http://www.inspire.anglia.ac.uk/e107_files/public/write_a_lot_better.pdf [Accessed 14 January 2011].

Ballantyne, R. (2004a) ‘Clydesdale gets backing from parent for English expansion’, *Sunday Times,* 5 September. pp. Business 3.

Ballantyne, R. (2004b) ‘Clyesdale to make £25m from new store’, *Sunday Times,* 12 September, pp. Business 3.

Barclays to sell 300 branches (2007) *Property Week,* 30 March, p13, London: United Business Media (UBM).

Barlas, Y. (1996) ‘Formal aspects of model validity and validation in system dynamics’, *System Dynamics Review,* 12(3), pp.183-210.

Barrie, G. (2006) ‘Bank on more good times for auctions’, *Property Week*, 4 August, p.33, London: CMP.

Bartlett, M. (1937) ‘Properties of sufficiency and statistical tests’, *Proceedings of the Royal Society of London Series,* 160, pp.268-24.

Baum, A. and Crosby, N. (1995) *Property Investment Appraisal,* London: Routledge.

Baum, A., Mackmin, D. and Nunnington, N. (2006) *The income approach to property valuation,* London: Reed Business Information (EG Books).

BBC invites building bids after Trillium departure (2005) *Property Week*, 9 September, London: CMP Information.

Beautifying bank branches (2001) *The Economist (US),* 24 March, p.6.

Beddy, M. (2010) ‘Lease is no longer more under new accounting regime’, *Property Week,* , pp.24, London: UBM Built Environment.

Bell, J. (2005) *Doing your research project: a guide for first-time researchers in education, health and social science*, Open University Press.

Benwell, J.B. (1771) ‘*An Essay on Interest and Annuities, embracing a concise investigation of the doctrine......’*, London: J.M. Richardson.

Berson, A., Smith, S. and Thearling, K. (2000) 'Building Data Mining Application for CRM', America: McGraw Hill Companies. In: 2000. Blacklick, Ohio: McGraw-Hill School Education Group, pp.90-105. Available through google.

Bids in for Germany's biggest deal (2005) 'Bids in for Germany's biggest deal', *Estates Gazette,*3 December, 548, p.36, London: Reed Business Information.

Bill, P. (2005) 'CBRE banks £30m outsourced deal', *Property Week*, 22 October, London: CMP Information.

Björklund, P. (2011) 'Hur påverkas hotellkoncerner av det nya förslaget för redovisning av leasingavtal?' [Swedish] *Externredovisning*, 16 November. Source: Google.

Blackhurst, C. (2008) 'Landlords are green if they think retailers will see reason', 6 September, *Estates Gazette*, p. 54, London: Reed Business Information.

Boots portfolio battle hots up as final four selected (2005) *Estates Gazette*, 14 May, p.31*.* London: Reed Business Information.

Boots to sell and lease back 300 stores (2005) *Property Week*. Available at http://www.property-week.co.uk/story.asp?storycode=3049352&sectioncode=297 [Accessed on 1 November, 2008].

Bourgeois, L. and Eisenhardt, K. (1988) 'Strategic decision processes in high velocity environments: Four cases in the microcomputer industry'. *Management Science,* pp.816-835.

Bowcock, P. (1978) *Property Valuation Tables,.* London and Basingstoke: MacMillan.

Bozdogan, H. (2000) 'Akaike's Information Criterion and Recent Developments in Information Complexity', *Journal of Mathematical Psychology,* 44, pp. 62-28.

Bradshaw, T. (2009) 'Tesco to open bank branches in stores', *The Financial Times,* 30 March, p.19, London: Pearson.

Bradshaw, T. (2009) 'Tesco to open in-store banks', *The Financial Times,* 30 March, p.19, London: Pearson.

Brasse, J. (2006) 'Philips pulls plug on £190m pension portfolio', *Property Week,* 29 September, p.7, London: CMP Information.

Brett, M. (2004) 'Unpacking property's Christmas presents', *Estates Gazette,* 18 December, p.36, London: Reed Business Information.

Britannia to sell surplus high street branches (2006) *Property Week,*17 February, p.7, London: CMP Information.

British Bankers' Association (2005a) *Main branded branch networks in Great Britain*, London: BBA Enterprises. Available at: http://www.bba.org.uk/jsp/polopoly.jsp?d=469&a=5524

British Bankers' Association (2005b) *Major banks' worldwide group assets at end of 2004*, p.1, London: BBA Enterprises. Available at: http://www.bba.org.uk/jsp/polopoly.jsp?d=469&a=5521

Bryman, A. (2001) *Social Research Methods.* 2nd ed. Oxford: Oxford University Press.

Burnyeat, A. (2005) ‘A round of business advice, if you please’, *Estates Gazette,* 26 March, pp.40-41, London: Reed Business Information.

Burton, D. (1990) ‘Competition in the UK retail financial service sector: some implications for the spatial distribution and function of bank branches’, *Service Industries Journal,* 10(3), pp. 571-588.

C & W takes in £64m (2006) *Estates Gazette*, 8 April, p.154, London: Reed Business Information.

Cahill, J. (2004a) ‘BBC questions Trillium contract’, *Estates Gazette*, 16 Oct, p.29, London: Reed Business Information.

Cahill, J. (2004) ‘Barclays puts its global portfolio under scrutiny’, *Estates Gazette*, 5 June, 423, p.31, London: Reed Business Information.

Cahill, J. (2004) ‘Departure casts pall over Mapeley’, *Estates Gazette,* 8 May, 419, p.29, London: Reed Business Information.

Cahill, J. (2004) ‘MPs quiz Mapeley and Revenue’, *Estates Gazette*, 30 October, p.39, London: Reed Business Information.

Cahill, J. (2004) ‘RBS jumps on outsourcing wagon’, *Estates Gazette,* 31 July, 431, p.21, London: Reed Business Information.

Cahill, J. (2004) ‘RBS to ask for bids in major outsourcing deal’, *Estates Gazette*, 20 November, p.31, London: Reed Business Information.

Cahill, J. (2005) ‘Ackerman makes RBS bid’, *Estates Gazette,* 2 April, p.29, London: Reed Business Information.

Cahill, J. (2005) ‘Bidders line up for £1.5bn NI outsourcing contracts’, *Estates Gazette,* 1 Oct, p.42, London: Reed Business Information.

Cahill, J. (2005) ‘Landsec Trillium banks Barclays outsourcing’, *Estates Gazette*, 15 Jan, p.28, London: Reed Business Information.

Cahill, J. and Lazarus, D. (2005) ‘BBC winds up Landsec Trillium outsourcing deal’, *Estates Gazette*, 14 May, p.30, London: Reed Business Information.

Cahill, J. and McClary, S. (2005) ‘Abbey banks on lure of leaseholds’, *Estates Gazette*, 7 May, P.41, London: Reed Business Information.

Cameron, S. (1997) *The MBA Handbook: Study Skills for Managers,* London: Financial Times Management.

Capita Symonds (2011) *The 2011 High Street lease renewal challenge,* London: Capita Symonds.

Cardew, B. (2004) ‘British Land completes £174m pub purchase’, *Property Week*, 20 October, London: CMP Information.

Carpenter, A. (2002) ‘Going their own way’, *Estates Gazette*, 9 March, 210, pp.51-54, London: Reed Business Information.

Catalano, A. (2002) ‘Securitisations set for 40% growth’, *Estates Gazette*, 202, 12 January, London: Reed Business Information.

Cavanagh, E. (2003) 'Assets on the Block', *Estates Gazette*, 22 February, pp.93-95, London: Estates Gazette Group.

Cavanagh, E. (2005) 'Sale-and-leaseback lots put on a show for CWHB', *Estates Gazette*, 12 March, p.133, London: Reed Business Information.

Cavanagh, E. (2005) 'Supermarket lots grab 4.8% at CWHB auction', *Estates Gazette,* 23 April, p.123, London: Reed Business Information.

Cavanagh, E. (2005) 'Harman Healy gets the cream with Devon lot', *Estates Gazette,* 1 October, p.113, London: Reed Business Information.

Cavanagh, E. (2005) 'Retail lots produce strong yields at Allsop auction', *Estates Gazette*, 15 October, p181, London: Reed Business Information.

Cavanagh, E. (2005) 'Leasebacks boost CWHB', *Estates Gazette*, 15 October, 541, p.182, London: Reed Business Information.

Cave, A. (2010) "Metro Bank: 'we'll put an end to stupid rules'", *The Daily Telegraph*, 20 February, London: Telegraph Media Group. Available at: http://www.telegraph.co.uk/finance/newsbysector/bankandfinance/7279848/Metro-Bank-well-put-an-and-to-stupid-rules.html

CB Richard Ellis (2008a) *UK Prime Rent & Yield Monitor.* London: CB Richard Ellis.

CB Richard Ellis (2008b) *UK Monthly Index.* London: CB Richard Ellis.

Chambers, J.M., Cleveland, W.S., Kleiner, B. and Tukey, P.A. (1983) 'Graphical methods for data analysis', Wadsworth & Brooks.

Check in to the hotel market (2005) 'Check in to the hotel market', *Property Week,* 26 August, p.99, London: CMP Information.

Chen, J., Hudson-Wilson, S. and Nordby, H. (2004) 'Real estate pricing: spreads & sensibilities: why real estate pricing is rational'. *Journal of Real Estate Portfolio Management,* 10(1), pp. 1-22.

Cheng, P. and Roulac, S.E. (2007) 'Measuring the effectiveness of geographical diversification', *Journal of Real Estate Portfolio Management,* 13(1), pp. 29.

Chesters, L. (2005a) 'Reit medicine for £300m Boots sale-and-leaseback', *Property Week*, 1 July, p.5, London: CMP.

Chesters, L. (2005b) 'Huge sell-off to follow Boots merger', *Property Week*, 7 October, p9, London: CMP Information.

Chesters, L. (2005c) 'Leslau gets Little Chefs for Christmas', *Property Week*, 16 December, p.1, London: CMP Information.

Chesters, L. (2006) 'HSBC to make 'mega' debut on Brompton Rd', *Property Week*, 27 October, p.9, London: CMP Information.

Clark, J. (1808) *An Enquiry into the Nature and Value of Leasehold Property, Reversionary Interests in Estates, and Life Annuities,* London: W. Walker.

Clark, S. (2005) 'What's all the fuss about opco-propco?', *Estates Gazette,* 5 May, p.59, London: Reed Business Information.

Coffer, A. (2002) 'Deutsche Bank eyes property cash', *Estates Gazette*, 28 September, 239, p.37, London: Reed Business Information.

Coffer, A. (2002) 'Bidders get their shouts in for £320 pub estate', *Estates Gazette*, 240, 5 October, p.39, London: Reed Business Information.

Coffer, A. (2004) 'Europe set for leaseback flood', *Estates Gazette*, 27 March, 413, p.26, London: Reed Business Information.

Collins, J. (1685) *The Doctrine of Decimal Arithmetick, Simple Interest, &c. As also of Compound Interest.* London: Printed by R. Holt for Nath. Ponder.

Collis, J. and and Hussey, R. (2003) *Business Research: A practical guide for undergraduate and postgraduate students,* 2nd ed. Basingstoke: Palgrave MacMillan.

Conover, W., Johnson, M.E., Johnson, M.M. (1981) 'A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data', *Technometrics,* 23, pp. 351-10.

Coolican, H. (1992) Research methods and statistics in psychology, p. 35, London: Hodder & Stoughton.

Corporate Finance, (2007) *Winning with branches.* Prague: Deloitte.

Creswell, J.W. (2009) *Research design: Qualitative, quantitative, and mixed methods approaches.* 3rd ed. London: Sage Publications, Inc.

Crosby, N. (2003) 'Accounting for leases', *Journal of Property Investment & Finance,* 21(2).

Crosier, S. (2007) 'Johann–Heinrich von Thünen: Balancing land-use allocation with transport cost', *Center for Spatially Integrated Social Science, University of California, Santa Barbara CA, USA, Available on:* http://www.csiss.org/classics/content/9 *[Accessed,* 25.

Cruise, S. (2005a) 'Not so S-mart', *Property Week*, 25 November, pp.32-35, London: CMP.

Cruise, S. (2005) 'Cheltenham & Gloucester offers £35m sale-and-leaseback', *Property Week*, 18 March, London: CMP. Available at: http://www.property-week.co.uk/story.asp?storycode=3048272&sectioncode=500 [Accessed 29 May 2005].

Cruise, S. (2005) 'Buyers warm to Iceland stores as Budget anxiety recedes', *Property Week*, 7 April, p.23, London: CMP.

Cruise, S. (2005) 'Helical cashes in on Unwins portfolio at Allsop auction', *Property Week*, 27 May, p.21, London: CMP Information.

Cruise, S. (2006a) 'Unwins' predicament leaves landlords in the lurch', *Property Week*, 6 January, p.17, London: CMP.

Cruise, S. (2006b) 'Khalastchi corrals £67m Barclays banks portfolio', *Property Week*, 29 September, p.13, London: CMP.

Cruise, S. (2006) 'Allsop bounces back with Budget Day sale', *Property Week*, 31 March, p.25, London: CMP.

Cruise, S. (2006) 'Buyers bank on Barclays portfolio at JLL's £82.1m sale', *Property Week*, 7 April, p.23, London: CMP.

Cruise, S. (2006) ‘Investors park bids for Kwik-Fit Portfolio’, *Property Week*, 8 December, p.7, London: CMP.

Cruise, S. (2006) ‘Retail regains buyers' confidence at Cushman sale’, *Property Week*, 7 April, London: CMP. Available at: http://www.property-week.co.uk/story.asp?storycode=3065525&sectioncode=500 [Accessed 11 May].

Cruise, S. (2007) ‘Barclays banks bump up Cushman sales...:Retail and bank sale and leasebacks tempt private investors’, *Property Week*, 9 March, p.51, London: CMP.

Cushman & Wakefield, (2008) *The Credit Crunch & Basel 2: Repairing the Balance Sheet with Real Estate,* April, London: Cushman & Wakefield.

Cushman & Wakefield, (2009) *Marketbeat: An Overview of the UK Property Market,* London: Cushman & Wakefield.

Cushman & Wakefield Healey & Baker (2004) *CWHB Auctions: auction results*, London: Cushman & Wakefield Healey & Baker. Available at: http://www.cushmanwakefieldeurope.com/auction

Cushman & Wakefield Healey & Baker, (2004) *Auction 25 February 2004,* catalogue, pp-.1-96, London: Cushman & Wakefield Healey & Baker.

Cushman & Wakefield Healey & Baker, (2004) ‘On the instructions of Bradford & Bingley plc for sale by auction’ (advertisement), *Estates Gazette,* London: Reed Business Information.

Cushman & Wakefield Healey & Baker, (2004) ‘Sale of 74 freehold, feuhold and leasehold bank, retail, office, industrial and leisure investments’ (advertisement), *Estates Gazette*, 17 February, p.15, London: Reed Business Information.

CWHB to advise Boots on sale-and-leasebacks (2005) ‘CWHB to advise Boots on sale-and-leasebacks’, *Estates Gazette,* 29 January, p.31, London: Reed Business Information.

D2 adds £72m HBOS pile to growing portfolio (2006) *Estates Gazette*, 5 August, p.31, London: Reed Business Information.

DB disposes of home branches (2004) *Euro Property*, p.3, ISSN 0967-1862, Business Premier Source.

Davey, J.(2005) ‘Property company with £1.4bn Revenue contract may float’. *The Times,* 02/12.

Davidson, A.W. (2002) *Parry's Valuation and Investment Tables.* London: Estates Gazette.

Dawe, T. and Thompson, H. (2005) ‘’It's the only game in town and this is how to play it’. *The Times,* 25 January, p. 54.

De Sautoy, O. (2009) *May 2009: UK Market Outlook.* London: Knight Frank.

Debenhams leaseback raises £495m (2005) *Estates Gazette,* 19 February, p.34, London: Reed Business Information.

Deloitte & Touche LLP (2007) *Winning with branches: six key disciplines for branch transformation.* London: Deloitte & Touche LLP.

Deutsche Bank (2004) *Deutsche Bank sells corporate real estate portfolio in Germany*, 27 December, Frankfurt am Main: Deutsche Bank. Available at:

http://www.db.com/presse/en/content/druckversion.htm [Accessed 7 September, 2006].

Deutsche Bank (2005a) *Biennes inmuebles: Toda la rentabilidad de una inversión sólida,* Barcelona: Deutche Bank, Sociedad Anónima Española.

Deutsche Bank (2005b) *Tu dinero llegará muy alto,* Barcelona: Deutche Bank, Sociedad Anónima Española.

Deutsche Bank (2005c) *Da en el blanco de la rentabilidad,* Barcelona: Deutche Bank, Sociedad Anónima Española.

Devaney, S. and Lizieri, C. (2004) 'Sale and leaseback, asset outsourcing and capital market impacts', *Journal of Corporate Real Estate,* 6(2), pp. 118-15.

Devlin, J. (1995) 'Technology and innovation in retail banking distribution', *International Journal of Bank Marketing,* 13(4), pp. 12-6.

Dorling, D. (2011) *'The North-South Divide - Where is the line?',* Available at: http://www.sasi.group.shef.ac.uk/maps/nsdivide/index.html [Accessed: 14 June].

Dover, M. (2005a) 'BT to engage adviser for £100m property cuts', *Property Week*, 7 October, p.11, London: CMP Information.

Dover, M. (2005b) 'Called to account', *Property Week*, 7 January, London: CMP Information. Available at: http://www.property-week.co.uk/story.asp?storycode=3045022&sectioncode=292

Dover, M. (2005c) 'LandSecs reasures over new accounting standards', *Property Week*, 11 February, London: CMP Information. Available at: http://www.property-week.co.uk/story.asp?storycode=3046642&sectioncode=37

Dover, M. (2005d) 'How will its garden centres grow?', *Property Week*, 8 July, pp.30-31, London: CMP Information.

Dover, M. (2005e) 'Wyevale to shake up property portfolio' *Property Week*, 13 July, London: CMP Information. Available at: http://www.property-week.co.uk/story.asp?storycode=3053780&sectioncode=297

Dover, M. (2005) 'Cable & Wireless axes property outsourcing plans', *Property Week*, 29 April, p.17, London: CMP Information.

Dover, M. (2005) 'Barclays slashes property advisers from 50 to 2', *Property Week*, 23 September, p.12, London: CMP Information.

Dover, M. (2005) 'Mapeley poised for £85m IBM sale-and-leaseback win', *Property Week*, 14 October, London: CMP Information. Available at: http://www.propertyweek.com/mapeley-poised-for-£85m-ibm-sale-and –leaseback-sin/3057655.article [Accessed 7 July 2011].

Dover, M. (2005) 'Slade's spin-off', *Property Week*, 11 November, London: CMP Information. Available at: http://www.propertyweek.com/slade's-spin-off/3058905.article [Accessed 7 July 2011].

Dover, M. (2006a) 'Royal Bank's second sale and leaseback: Bank to market second UK-wide portfolio', *Property Week*, 2 June, p.9, London: CMP Information.

Dover, M. (2006b) 'Banks to the future', *Property Week,* 4 August, pp.28-30, London: CMP Information. Available at: http://www.property-week.com/banks-to-the-future/3071741.article [Accessed 7 July 2011].

Dover, M. (2006) 'Mapeley's passport to outsourcing comeback', *Property Week*, *27 January,* London: CMP Information. Available at: http://www.propertywek.com/mapeleys-passport-to-outsourcing-comeback/3061804.article [Accessed 7 July 2011].

Dover, M. (2006) 'Property shake-up at Royal Bank of Scotland', *Property Week*, 24 February, p.9, London: CMP Information.

Dover, M. (2006) 'Sale and leasebacks are more popular than ever', *Property Week*, 17 March, p.23, London: CMP Information.

Dover, M. (2006) 'RBS appoints Cushman in global property shake-up', *Property Week*, 7 April, p.13, London: CMP Information.

Dover, M. (2006) 'Barclays hands over property strategy to managers', *Property Week*, 28 April, p.14, London: CMP Information.

Dover, M. (2006c) 'HSBC designs £400m return to high street banking', *Property Week,* 30 June 2006, p. 9, London: CMP Information.

Drake, P. and Heath, L. (2011) *Practitioner Research at Doctoral Level,* 1st ed. Oxford: Routledge.

Dunse, N., Jones. C., White, M., Trevillion, E. and Wang, L. (2007) 'Modelling Urban Commercial Property Yields: Exogenous and Endogenous Influences', *Journal of Property Research,* 24(4), pp.335-20.

Duxbury, R. (2010) 'Banks still the safest place to put your money', *Property Week,* 76(4), p.87. London: UBM Built Environment.

Eade, C. (2006) 'Don't bank on it', *Property Week*, 4 August, pp.62-63, London: CMP.

Edelstein, H. (1996) 'Data Mining: Exploiting the Hidden Trends in Your Data', *DB2 Online Magazine,* Available through: Google.

Edwards, T. (2005) 'Crofting', *Scottish Parliament Information Centre briefing,* 05/50, pp.1-15.

Efron, B. and Gong, G. (1983), 'A leisurely look at the bootstrap, the jackknife, and cross-validation', *American Statistician,* pp.36-48.

Efron, B. and Tibshirani, R. (1997) 'Improvements on cross-validation: The. 632 bootstrap method', *Journal of the American Statistical Association,* pp.548-560.

Efron, B. and Tibshirani, R. (1986) 'Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy', *Statistical science,* pp.54-75.

Enever, N. and Isaac, D. (2002) *The valuation of property investments.* London: Estates Gazette Ltd.

Enever, N. (1984) *The Valuation of Property Investments.* London: Estates Gazette.

Evans, A.W. (1995) 'The property market: ninety per cent efficient?', *Urban Studies,* 32(1), pp.5-30.

Every little deposit helps: Tesco goes into banking (2009) *Sunday Times,* 29 March, Features p.7, London: Times Newspapers.

Fagan, J, (2006) 'Bank to sell and lease back 36 branches', *The Irish Times*, 26 July. Available at: http://www.ireland.com/newspaper/commercialproperty/2006/0726/pf115381377878 [Accessed 7 September 2006].

Fawcett, T. (2006) 'An introduction to ROC analysis', *Pattern Recognition Letters,* 27, pp.861-874.

Field, A.P. (2005) *Discovering Statistics Using SPSS,* 2nd ed. London: Sage Publications Inc.

Fisher, J. (1996) *Mathematics For Real Estate Appraisers,* Illinois: Appraisal Institute.

Fisher, W.E. (1960) National certificate mathematics, v. 1, London: The English Universities Press Ltd, pp.131-158.

Fitz-Gibbon, C.T. and Morris, L.L. (1987) *How to Analyze Data,* Newbury Park: Sage Publications.

Flanagan, P. (2006) 'After 125 Years, the postal order gets a new look', *Daily Express,* 27 April, p.33, London: Express Newspapers.

Flynn, T. (1997) 'The supermarket branch revisited: how good are they, really, when you crunch the numbers?', *ABA Banking Journal,* 89(10), pp.24-26.

Fornero, R. (2007) *Cronología Fotográfica de las Finanzas: Los instrumentos, los cenceptos, las herramientas,* Mendoza, Argentina: Universidad Nacional de Cuyo.

Foxley, S. (2005) 'In search of a rational property market', *Estates Gazette*, 29 January, 504, p.54, London: Reed Business Information.

Franks, C. (2005) 'Boots to sell and lease back 300 stores', *Property Week*, 7 April, London: CMP. Available at: http://www.propertyweek.com/news/boots-to-sell-and-lease-back-300-stores/3049352.article [Accessed 7 July 2011].

Fraser, A. (2003) 'Branch increase leads bank to profit', *Australian Business Intelligence,* September 2004, pp. 1008094i1493.

Fraser, W.D. (2004) *Cash-flow Appraisal for Property Investment,* Basingstoke: Palgrave.

French, N. (2004) 'The valuation of specialised property: A review of valuation methods', *Journal of Property Investment and Finance,* 22(6), pp.533-9.

Freshfields Bruckhaus Deringer (2004) *Deutsche Bank completes €300m sale-and-leaseback of corporate real estate portfolio*, 30 December 2004, p.1-2, Freshfields Bruckhaus Deringer. Available at: http://www.freshfields.com/new/dynamic/Pressrelease.asp?newsitem=599 [Accessed 7 September 2006].

Gallimore, P., Hansz, J.A. and Gray, A. (2000) 'Decision making in small property companies', *Journal of Property Investment and Finance,* 18(6), pp.602-11.

Gentle, C. (2008) 'Retail is in the detail: how banks need to re-energize retail franchises', *The Journal of Risk Management,* 9(1).

George, J. and Pazzi-Axworthy, R. (2002) 'Outsource your resource', *Estates Gazette*, 23 March, pp.128-130, London: Reed Business Information.

German fund buys Barclays' card HQ (2006) 'German fund buys Barclays' card HQ', *Estates Gazette*, 17 June, p.30, London: Reed Business Information.

Gibson, R. (2006) 'Blackstone floats £800m plans to buy back villages', *Estates Gazette*, 11 March, p.53, London: Reed Business Information.

Gibson, R. (2005) 'Slice of Iceland deal up for grabs', *Estates Gazette*, 15 January, p.33, London: Reed Business Information.

Gilbertson, B. (2009) 'Leaderboard', *RICS Business,* February, pp.7, London: Royal Institution of Chartered Surveyors.

Gordon, W. (2005) 'There's no danger of VAT', *Property Week*, 3 June, London: CMP. Web: http://www.property-week.co.uk/story.asp?storycode=3051781&sectioncode=274

Grand Hotels (2005) *Property Week,* 22 July, London: CMP Information. Web: http://www.property-week.co.uk/story.asp?storycode=3054036&sectioncode=286

Greenland, S. (1995) 'Network management and the branch distribution channel', *International Journal of Bank Marketing,* 13(4), pp.12-18.

Grierson, J., Davis, R. and French, N. (2006) 'Let go with control', *Estates Gazette*, 18 February, pp.130-132, London: Reed Business Information.

Guy, S. and Henneberry, J. (2000) 'Understanding urban development processes: integrating the economic and the social in property research', *Urban Studies,* 37(13), pp.2399.

Hall, J. (2009) '*Tesco to launch bank branches and current accounts',(City),* London: Telegraph Media Group. Web: http://find.galegroup.com/itx/infomark.do?&contentSet=IAC-Documents&type=retrieve&tabID=T004&prodId=SPN.SP00&docId=CJ196596144&source=gale&srcprod=SP00&userGroupName=anglia_itw&version=1.0

Harman Healy takes in £15m (2004) 'Harman Healy takes in £15m', *Estates Gazette*, 17 December, p.75, London: Reed Business Information.

HBOS (2008) *Recommended Acquisition of HBOS plc by Lloyds TSB Group plc to be effected by means of a scheme under sections 895 to 899 of the Companies Act 2006: Circular to HBOS Shareholders and Notice of HBOS Court Meeting and General Meeting,* Edinburgh: HBOS plc.

Healey & Baker (1997) *On behalf of Midland Bank 99 Freehold and Long Leasehold Bank Investments (catalogue),* pp.1-144, London: Healey & Baker.

Hidden jewels: bank branches (1993) *The Economist (US),* vol.62(2), pp.171-181.

Hipwell, D. (2007) 'Bank of Ireland branches out for €100m', *Property Week*, 20 April, p.17, London: CMP.

Hipwell, D. (2006) 'Banking the profits', *Property Week*, 8 September, pp.9-10, London: CMP.

H.M. Revenue & Customs (2005) 'SDLTM10000 Leases', *Stamp Duty Land Tax Manual*, pp.32-33, London: HMRC. Available at: http://www.hmrc.gov.uk/so/sdlt-lease.pdf [Accessed 6 October 2006].

H.M. Revenue & Customs (2010a) *Stamp Duty Land Tax rates and thresholds*, London: HMRC. Available at: http://www.hmrc.gov.uk/sdlt/rates-thresholds.htm [Accessed 3 April 2010].

H.M. Revenue & Customs (2010b) 'Disadvantaged Areas Stamp Duty Exemption, *Customer Newsletter*, London: HMRC. Available at: http://www.hmrc.gov.uk/SO [Accessed 3 April 2010].

Ho, G. (2011) 'Banking on a new approach: Metro's chairman is happy with progress', *Sunday Express,* 10 July, pp. Financial 2, London: Express Newspapers.

Hoesli, M., Lizieri, C. and MacGregor, B. (1997) 'The spatial dimensions of the investment performance of UK commercial property', *Urban Studies,* 34(9), pp.1475.

Hoesli, M. and MacGregor, B.D. (2000) *Property investment: principles and practice of portfolio management,* Pearson Education.

Holmes, D., Moody, P. and Dine, D. (2006) *Research methods for the biosciences*, p.177, Oxford: Oxford University Press.

HSBC Holdings plc (2006) 'HSBC Target Towns', *Estates Gazette*, 1 July, advertisement between pp. 24 and 25, London: Reed Business Information.

Imber, A. (2005) 'From owner to tenant', *Estates Gazette*, 3 September, 535 pp.108-109, London: Reed Business Information.

Investment Property Databank (2009) *IPD UK Annual Property Index 2008,* London: Investment Property Databank.

Investment Property Databank (2008) *IPD United Kingdom retail property index,* London: Investment Property Databank.

Investors lack faith in Wales (2005) 'Investors lack faith in Wales', *Estates Gazette*, 25 June, p.35, London: Reed Business Information.

Inwood, J. (1880) *Tables for the Purchasing of Estates, Freehold, Copyhold, or Leasehold; Annuities, Advowsons, &c.,* London: Crosby, Lockwood & Co.

Isaac, D. and Steley, T. (2000) *Property Valuation Techniques.* Basingstoke: Palgrave.

Jenkins, S. (2002) ''Mob Scene' awaits H & B', *Estates Gazette*, 16 February, 207, p.92, London: Reed Business Information.

Jenkins, S. (2004) 'Pubs and surgeries power Allsop to £75m auction', *Estates Gazette*, 11 December, 450, p.63, London: Reed Business Information.

Jones Lang LaSalle (2007) *The Private Investor in the UK High Street,* London: Jones Lang LaSalle.

Jones Lang LaSalle (2008a) *Pulse: UK Property Index Q1 2008,* London: Jones Lang LaSalle.

Jones Lang LaSalle (2008b) *On Point: In Search of Value? What market conditions are office occupiers facing?* London: Jones Lang LaSalle.

Jones Lang LaSalle (2009 a) *Pulse: UK Property Index Q1 2009,* London: Jones Lang LaSalle.

Jones Lang LaSalle (2009 b) *On Point: Auction Results Analysis System Report - Q1 2009.* London: Jones Lang LaSalle.

Jones, S. (2003) 'Bank branch openings hit mortgage brokers', *Australian Business Intelligence,* 3 April, pp.1008093i0823.

Kane, E. (1987) '*Doing Your Own Research: Basic Descriptive Research in the Social Sciences and Humanities',* London: Marion Boyars Publishers.

Kane, E. and O'Reilly-De Brün, M. (2005) *Doing Your Own Research.* London: Marion Boyars.

Kao, B. and Zhang, M. (2003) 'Algorithms for mining frequent sequences', *Advances in Data Mining and Modeling,* p.1, New Jersey: World Scientific Publishing.

Kelleher, E. (2004), 'Banking on a branch line', *The Financial Times,* 22 September, p.6, London: Pearson.

Kelly, J. (2010) 'The High Street Bank re-imagined', *BBC News Magazine*, 14 July, London: BBC New. Available at: www.bbc.co.uk [Accessed 29 March 2011].

Khalastchi to sell £12.5m portfolio (2005) 'Khalastchi to sell £12.5m portfolio', *Property Week,* London: CMP Information.

King Sturge (2005) 'Former bank premises for sale', *Estates Gazette,* 22 October, p.21, London: Reed Business Information.

Kingsmill, R. (2005) 'The journey from owner to occupier', *Estates Gazette,* 6 August, 531, pp.74-76, London: Reed Business Information.

Kirk, J. and Miller, M.L. (1986) *Reliability and validity in qualitative research,* Sage.

Kivlehan, N.P. (2011) 'Calling the tune', *Estates Gazette,* 21 May, pp.83-85, London: Reed Business Information.

Kleinman, M. (2005) 'CVC revs up for Kwik-Fit sale', *The Sunday Times,* 20 February, p. 3.2, London: Times Newspapers.

Klenowski, V. and Lunt, I. (2008) 'Enhancing Learning at Doctoral Level through the Use of Reflection?', *Assessment & Evaluation in Higher Education,* 33(2), pp.15.

Knight Frank (2009) *UK Market Outlook,* London: Knight Frank.

Knight, P.T. (2002) *Small-Scale Research:Pragmatic Inquiry in Social Science and the Caring Professions.* London: Sage Publications.

Kohavi, R. (1995) 'A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection', *International Joint Conference on Artificial Intelligence*, pp.1137-43.

KPMG (2006) *Sale and leaseback takes the spotlight,* London: KMPG LLP.

Krystalogianni, A. and Tsolacos, S. (2004) 'Regime switching in yield structures and real estate investment', *Journal of Property Research,* 21(4), p.279.

Kuhn, T.S. (1970) 'The structure of scientific revolutions', *Chicago, IL*.

Lambert Smith Hampton Research (2009) *UKIT Quarterly Bulletin 2009 Q1,* London: Lambert Smith Hampton.

LandSecurities Trillium (2004) ‘Outsourcing’, *LandSecurities Trillium*. Web: http://www.lstrillium.co.uk

La Poste (2007) *La Banque Postale, bien plus qu’une banque*, Paris: La Poste. Available at: http://www.labanquepostale.fr [Accessed 31 May 2007].

Lawson, D. (2001) ‘Property's balancing act’, *Property Week*, 13 August, London: CMP. Available at: http://www.propertyweek.com/propertys-balancing-act/3008174.article [Accessed 7 July 2011].

Lawson, D. (2004) ‘EU holds property to account’, *Property Week,* 1 March, London: CMP. Available at: http://www.propertyweek.com/eu-holds-property-to-account/3039349.article [Accessed 7 July 2011].

Lazarus, D. and Cahill, J. (2004) ‘Bank switches from let to buy for £120m City HQ’, *Estates Gazette,* 6 November, 445, p.30, London: Reed Business Information.

Lee, S. and Byrne, P. (1998) ‘Diversification by sector, region or function?’, *Journal of Property Valuation and Investment,* pp.38-19.

Leedy, P.D. and Ormrod, J.E. (2010) *Practical research: Planning and design,* New Jersey: Pearson Education.

Leshem, S. and Trafford, V.N. (2007) ‘Overlooking the conceptual framework’, *Innovations in education and Teaching International,* 44(1), pp.93-105.

Levin, G. (2004) ‘Is the Time Right for a Sale-Leaseback?’, *Finance Executive*, September, 20(6), pp.58-59.

Linder, J. (2004) ‘Outsourcing as a strategy for driving transformation’, *Strategy and Leadership,* 32(6), pp.26-6.

Locke, D. (2006) ‘*A reaction for every action’,* Estates Gazette, 6 June, pp.142-144, London: Reed Business Information.

Louargand, M. (1992) ‘A Survey of Pension Fund Real Estate Portfolio Risk Management Practices’, *Journal of Real Estate Research,* pp.361-13.

Macgregor, B. and Schwann, G. (2003) ‘Common features in UK commercial real estate returns’, *Journal of Property Research,* 20(1), pp.23-48.

Mackmin, D. (2009) ‘Put the bases to the test’, *Estates Gazette,* 26 September, pp.106-8, London: Reed Business Information.

Mapeley Limited (2004) ‘Client overview: Abbey’, Mapeley. Available at http://www.mapeley.com/abbey.asp (Accessed: 1 November 2008).

Mapeley takes STEPS to refinance £176m (2006) *Estates Gazette*, London: Reed Business Information.

Markowitz, H. (1952) ‘Portfolio selection’, *Journal of Finance,* pp.77-91.

Markowitz, H.M. (1983) ‘Nonnegative or not nonnegative: A question about CAPMs’, *Journal of Finance,* 38(2), pp.283-13.

Marriott, R. (2005) ‘Hunter and Leslau plan Big Food sale-and-leasehold’, *Property Week,* 7 January, London: CMP. Available at: http://www.propertyweek.com/news/hunter-and-leslau-plan-big-food-sale-and-leaseback/3045039.article [Accessed 7 July 2011].

Marshall, P. (1988) *Donaldsons' Investment Tables,* London: Donaldsons Chartered Surveyors.

Mathison, R. (2001) 'Counter Attack'. *Scottish Banker,* p.20, Edinburgh: Chartered Institute of Bankers in Scotland.

Maxwell, E. (2006a) 'Investors seek long-term security in auction rooms', *Estates Gazette*, 19 April, p.93, London: Reed Business Information.

Maxwell, E. (2006b) 'Blue-chip multiples flock to sale and leaseback', *Estates Gazette*, 24 June, p.94, London: Reed Business Information.

Maxwell, E. (2006c) 'Totals down at C&W July sale', *Estates Gazette,* 15 July, p.66, London: Reed Business Information.

Maxwell, E. (2006d) 'Sales rate dips below 80%', *Estates Gazette*, 22 July, p.87, London: Reed Business Information.

Maxwell, E. (2006e) '£47.6m taken at C&W', *Estates Gazette,* 9 December, p.69, London: Reed Business Information.

Maxwell, J. (1992) 'Understanding and Validity in Qualitative Research', *Harvard educational review,* 62(3), pp.279-300, Harvard University.

Maxwell, J.A. (2002) 'Understanding and validity in qualitative research', *The qualitative researcher's companion,* pp.37-64.

McKenna, C. (2005) 'Unforeseen dangers of sale and manage back contracts', *Royal Institution of Chartered Surveyors website*, p.1-2, London: Royal Institution of Chartered Surveyors. Available at: http://www.rics.org/Property/Commercialproperty/salemanageback%20contracts.html [Accessed 27 May 2005].

McClary, S. (2005a) 'Investors kick Boots portfolio offer', *Estates Gazette*, 16 April, p.29, London: Reed Business Information.

McClary, S. (2005b) 'Tesco shops for asset management role', *Estates Gazette*, 24 September, p.43, London: Reed Business Information.

McClary, S. (2006) 'Don't think Big Mac, think McProperty', *Estates Gazette*, 11 March, pp.66-69, London: Reed Business Information.

McClary, S. (2010) 'Rules change may lead to shorter lease terms', *Property Week,* 21 August, p.25, London UBM Built Environment.

McKenna, C. (2005) 'Unforeseen dangers of sale and manageback contracts', pp.1-2, Available at: http://www.rics.org/Property/Commercialproperty/salemanageback%20contracts.html [Accessed: 27 May, 2005].

McMillan, R. (2011) 'Taking a new approach to financial statements', *Commercial Property Journal,* May-June, pp.12-13, London: Royal Institution of Chartered Surveyors.

Merrell, C. (2004) 'Alliance & Leicester to close 46 branches', *The Times,* 17 June, London: Times Newspapers.

Metro Bank (2011) *A Revolution in Banking*, pp.1-2, London: Metro Bank plc. Available: https://metrobankonline.co.uk (Accessed: 29 March, 2011).

Metro banks two more locations (2010) 'Metro banks two more locations', *Estates Gazette*, 3 April, p.16, London: Reed Business Information.

Moon, J.A. (1999) *Learning Journals: A Handbook for Academics, Students and Professional Development.* London: Kogan Page.

Moore, N. (2000) *How To Do Research: The Complete Guide to Designing and Managing Research Projects.* London: Library Association Publishing.

Morgan, G. and Smircich, L. (1980) 'The case for qualitative research', *Academy of management review,* 5(4), pp.491-500.

Morland, Sir Samuel (1679) *The Doctrine of Interest, ... explained, ... discovering the erors of the ordinary Tables of Reb..* London.

Morrison, J. (2005) 'Grand hotels', *Property Week*, 22 July, pp.67-68, London: CMP Information.

Mortished, C. (2004) 'UK's first Islamic bank confirms £40m float', *The Times,* 27 August, p. 4, London: Times Newspapers.

Mychasuk, E. and Terazono, E. (2008) 'Tesco bank balance tipped', *The Financial Times*, 2 October, p.20, London: Pearson.

National Australia seeks adviser for UK estate (2005) 'National Australia seeks adviser for UK estate', *Property Week*, 11 March, London: CMP Information. Available: http://www.property-week.co.k/story.asp?storycode=3048052&sectioncode=36 [Accessed 2 August 2008].

Nellis, J. and Lockhart, T. (1995) 'The impact of deregulation on the UK building society branch network in the 1990s', *International Journal of Bank Marketing,* 13(4), pp.5-11.

Nelson, T.R. and Nelson, S.L. (2003) 'Regional models for portfolio diversification', *Journal of Real Estate Portfolio Management,* 9(1), pp.71-88.

Newton, Sir Isaac (1731) *A true Estimate of the Value of Leasehold Estates, and of Annuities and Reversions for lives and..,* London.

Nisse, J. (2004) 'HBOS calls in KPMG integrators to make £10bn Abbey bid add up', *The Independent on Sunday,* 12 September, p. Business 1, London: Independent Newspapers.

Norman, P. (2005) 'Lots to watch out for', *Estates Gazette*, 26 February, pp.165-166, London: Reed Business Information.

Norman, P. (2005) 'Sale-and-leasebacks take star place in MOD arsenal', *Estates Gazette*, 28 May, p.33, London: Reed Business Information.

Northedge, R. (2005) 'Time to split', *Estates Gazette*, 5 March, pp.71-75, London: Reed Business Information.

Nurnberg, H. and Largay, J. (1996) 'More Concerns Over Cash Flow Reporting Under FASB Statement No. 95', *Accounting Horizons,* pp.123-7.

O'Connell, B. (1998) 'Banks and supermarkets - now for the big sell', *Australian Banker,* 112(5), pp.192-194, Gale.

O'Connor, J. (2005) 'UK companies unlock wealth from their bricks and mortar', *The Lawyer*, 19(8), pp.8-9.

O'Grady, S. (2006) 'Post offices to install free cash machines', *Daily Express,* 23 October. p. 27, London: Express Newspapers.

Omar, A.J. and Heywood, C. (2010) 'Corporate real estate management's credibility-positioning status: Preliminary investigation of a branding model in practice', *Journal of Corporate Real Estate,* 12(3), pp.185-11.

O'Neill, M. and Mathews, K. (2000) 'A Weighted Least Squares Aroach to Levene's Test of Homogeneity of Variance', *Australian & New Zealand Journal of Statistics,* 42(1), pp.Abstract. Available .

Orna, E. and Stevens, G. (2000) *Managing Information for Research.* Buckingham: Open University Press.

Our post offices will still need protecting (2006) *The Daily Telegraph,* 14 December, p. 26, London: Telegraph Media Group.

Outsourcing 'poor value' (2005) 'Outsourcing "poor value"', *Estates Gazette*, 19 March, p.35, London: Reed Business Information.

Owen, F. and Jones, R. (1994) *Statistics,* 4th ed. Harlow (England): Pearson Education.

Paisley, R. (2000) *Green's concise Scots law,* Edinburgh: Green & Sons/Sweet & Maxwell.

Parnham, P. and Rispin, C. (2005) *Home Inspector's Handbook,* May 2005, pp.96-97, Coventry (UK): RICS Books.

Passport stamp of approval for Mapeley , 2006. *Passport stamp of approval for Mapeley.* London: CMP Information.

Pension problems will drive leaseback boom (2006) 'Pension problems will drive leaseback boom', *Estates Gazette*, 18 March, p.45, London: Reed Business Information.

Phillips, M. (2007) 'Banks set Cushman sale to good account', *Property Week*, 25 May, p.31, London: CMP Information.

Pinheiro, J. and Bates, D. (2009) *Akaike's An Information Criterion,* Bristol: CRAN.

Posada, D. and Buckley, T. (2004) 'Model Selection and Model Averaging in Phylogenetics: Adavantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests', *Systemic Biology,* 53(5), pp.793-15.

Post Office (2003) *A little short? Take out cash here free,* London: Post Office Limited.

Post Office (2005) *Get your loan from a name you know,* London: Post Office Limited.

Post Office futures to be announced (2006) *Global Internet News*. Available at: http://www.global.net.uk/news/news.asp?aid=18009576&cat=news [Accessed 14 December 2006].

Pottinger, G., Dixon, T. and Marston, A. (2002) 'Occupational futures? Divesting real estate and corporate PFI', *Property Management,* 20(1), pp.31-48.

Poulter, S. (2004) 'The great free cash machine betrayal', *Daily Mail,* 16 November, pp.1-2

Poulter, S. (2006) 'Banks open branches - to sell loans and credit cards', *Daily Mail,* 6 June, p. 26.

Pratt, A. (2008) 'Counter Revolution', *RICS Business,* Nov/Dec, pp.26-27, London: Royal Institution of Chartered Surveyors.

Prendergast, G. and Marr, N. (1994) 'Towards a Branchless Banking Society?' *International Journal of Retail & Distribution Management,* 22(2), pp.18-26.

Preston, D. (2004) 'Christmas present is great value: Accounting standards add dimension to professional advice', *Property Week*,17 December, London: CMP Information. Available at: http://www.propertyweek.com/professional/christmas-present-is-great-value/3044634.article [Accessed 7 July 2011].

Price, R. (1771) *Observations on Reversionary Payments; on schemes for providing annuities for widows ... and on t.* London.

Punch, K.F. (2003) *Survey Research: The Basics.* London: Sage Publications.

R Development Core Team (2010) ISBN 3-900051-07-0, URL http://www.R-project. org.

Ranjan, J. and Malik, K. (2007) 'Effective educational process: a data-mining approach', *VINE - The Journal of Information and Knowledge Management Systems,* 37(4), pp.502.

Ratcliffe, A. (2005) 'The well-filled shopping cart', *Estates Gazette*, 29 October, pp.180-182, London: Reed Business Information.

RBS in innovative sale and leaseback (2004) 'RBS in innovative sale and leaseback', *Estates Gazette,* 30 October, p.37, London: Reed Business Information.

Reuveni, G. (2002) 'Reading sites as sights for reading. The sale of newspapers in Germany before 1933: bookshops in railway stations, kiosks and street vendors', *Social History,* 27(3), pp.273-15.

Ricardo, D. (1815) *An Essay on Profits.* London: Murray.

Romney, C. (2005) 'Beware the hidden perils of the management agreement', *Property Week*, 20 May, p.79, London: CMP Information.

Rose, J. (1993) *Square Feet,* London: RICS Books.

Rose, J.J. (1975) *Property Valuation Tables,* Oxford: The Freeland Press.

Rouse, W. (1816) *An Investigation of the errors of all writers on annuities in their valuation of half-yearly and quarterly payments,* London: Lackington, Allen & Co.

Rowe, S. and Cahill, J. (2005) 'VAT victory for Abbey outsourcing', *Estates Gazette*, 14 May, p.34, London: Reed Business Information.

Royal Institution of Chartered Surveyors (2011) *The New Lease Accounting Rules: An overview of the proposals' impact on Corporate Real Estate,* London: Royal Institution of Chartered Surveyors.

Royston, P. (1995) 'Remark AS R94: A remark on Algorithm AS 181: The W test for normality', *Applied Statistics,* 44, pp.547-4.

Russell, J. (2007) 'Mapeley to hit acquisition target', *Estates Gazette*, 4 November, 644, p.55, London: Reed Business Information.

Sandher, H. (2010) 'The banks that like to say yes', *Property Week*, 16 July, pp.30-33, London: UBM Built Environment.

Savings Watch: The Post Office (2005) *The Times,* 24 September, pp. Money 6, London: Times Newspapers.

Schafer, S. (1994) 'Bank branch valuation: an empirical approach', *Appraisal Journal,* 62(2), pp.171-180.

Schäfer-Surén, J. (2005) 'Make the right choice from hotel lease menu', *Estates Gazette*, 3 December, p.48, London: Reed Business Information.

Schön, D. (2003) *The Reflective Practitioner: How Professionals Think In Action,* Aldershot (UK): Ashgate.

Scorgie, M.E. (1996) 'Evolution of the Application of Present Value to Valuation of Non-monetary Resources', *Accounting and Business Research,* 26(3), pp.237-11.

Senior-Milne, G. (2004) 'Farewell to feudalism', *The Scottish Genealogist*, November. Available: *http://www.scotgenealogy.com/online/farewell_to_feudalism.htm*

Shah, A. (2010) 'Fear of change as occupiers account for losses', *Property Week,* 3 September, p.52, London UBM Built Environment.

Shepherd, M. (2006) 'Barclays banks on Denton for property', *Property Week*, 28 July, p.61, London: CMP Information.

Simpson, T. (1742) *The Doctrines of Annuities and Reversions, ... with ... tables, etc.* London.

Sing, T., Sander, O., Beerenwinkel, N. And Lengauer, T. (2009) 'ROCR: Visualizing the performance of scoring classifiers', *CRAN R-project*, pp.1-3, Bristol: CRAN. Available at: http://www.CRAN.R-project.org/package=ROCR.R_package_version [Accessed 14 November 2010].

Spooner, S. (2005) 'Principal players', *Estates Gazette*, 26 November, pp.140-141, London: Reed Business Information.

Stamp Duty: The Post Office network can never be absolutely cost efficient (2006) *The Times,* 15 December 2006 p.20, London: Times Newspapers.

Stanfield, M. (2004) 'Outsourcing thoughts turn to Europe Investment Climate', *The Financial Times,* 22 September, p. 11, London: Pearson.

Steinborn, D. (1994) 'Time to rethink in-store banking', *ABA Banking Journal,* 86(11), pp.43-3.

Supermarkets fail to make it a clean sweep (2005) 'Supermarkets fail to make it a clean sweep', *The Times,* 24 September, p. Money 6, London: Times Newspapers.

Tabachnick, B. and Fidell, L. (2001) *Using Multivariate Statistics.* 5th ed. Needham Heights, MA, USA: Allyn & Bacon.

Tesco bank plays its reward card (2008) *The Sunday Times*, 16 November, p.7(Business), London: Times Newspapers.

Thame, D. (2010) ‘Every little deal helps’, *Estates Gazette*, 11 September, pp.87-89, London: Reed Business Information.

The Banker, (2004) *Retail Strategy: Strategy - UK Post Offices Sees Market For 'No Frills' Financial Services.* London: Financial Times Ltd.

The golden branch (1992) *The Economist (US),* 2 May, pp.542-543.

The misery-go-round (2005) ‘The misery-go-round’, *The Mail on Sunday Personal Finance,* 4 September, p. 9, London: Associated Newspapers.

The Scottish Parliament Information Centre (2000) *Research Paper 00/09: The Abolition of Feudal Tenure etc. (Scotland) Bill.*

Thomas, D. (2006a) ‘Barclays banks on property sell-off with £250m package’, *Property Week*, 24 March, p.14, London: CMP Information.

Thomas, D. (2006b). ‘Prudential procures £85m high street Barclays portfolio’, *Property Week*, 5 May, p.12, London: CMP Information.

Thomas, R. (1823) *Tables of the values of Estates, Annuities, Assurances on lives, etc.* Falmouth: R. Thomas.

Tinworth, A. (2005a) ‘How we got here’, *Estates Gazette*, 24 September, pp.76-77, London: Reed Business Information.

Tinworth, A. (2005b) ‘In search of an all-seeing eye’, *Estates Gazette*, 1 October, pp.70-71, London: Reed Business Information.

Tipping, M. (2006) *The Impact of the Time Value of Money on Valuation Practice,* XXIII FIG Congress Munich, pp.1-15, Copenhagen: FIG.

Tipping, M. and Bullard, R. (2007) *Sale-and-leaseback as a British Real Estate Model,* XXX FIG Working Week 2007, Hong Kong, pp.1-15, Copenhagen: FIG.

Tipping, M. and Bullard, R.K. (2007) ‘Sale-and-leaseback as a British real estate model’, *Journal of Corporate Real Estate,* 9(4), pp.205-17.

Tipping, M. and Lam, T. (2010) *Factors most influencing United Kingdom retail property yields: a theoretical perspective,* COBRA Conference, Paris, 2nd-3rd September, 2010, London: Royal Institution of Chartered Surveyors.

Trafford, V. and Leshem, S. (2008) *Stepping stones to achieving your doctorate,* Maidenhead: Mcgraw Hill/Open University Press.

Tyler, R. (2009) ‘Tesco launches bank branches in its stores’, *The Daily Telegraph* (*City),* 30 March, London: Telegraph Media Group. Available at http://find.galegroup.com/itx/infomark.do?&contentSet=IAC-Documents&type=retrieve&tabID=T004&prodId=SPN.SP00&docId=CJ196668502&source=gale&srcprod=SP00&userGroupName=anglia_itw&version=1.0

Under fire: Europe's postal banks (1992) *The Economist (US),* 20 June, pp.81-82.

Upward-only rents cripple Southern Cross (2011) *Estates Gazette*, 4 June, p.40, London: Reed Business Information.

Vernham, P. (2011) ‘Lease said, soonest mended’, *Estates Gazette,* 5 March, p.87, London: Reed Business Information.

Von Thünen, J.H. and Hall, P.G. (1966) *Isolated state: an English edition of Der isolierte Staat,* Pergamon Press.

Walker, G. (2009) *Allsop release their July commercial auction catalogue.* Available at: http://www.allsop.co.uk [Accessed: 2009].

Webb, J. (1984) 'Real Estate Investment Acquisition Rules for Life Assurance Companies and Pension Funds: A survey', *AREUEA Journal,* 12(4), pp.495-520.

Weiss, S.M. and Indurkhya, N. (1998) *Predictive Data Mining: A Practical Guide.* Morgan Kaufmann.

West LB seeks to cash in on freeholds (2006) 'West LB seeks to cash in on freeholds', *Estates Gazette,* 18 March, p.32, London: Reed Business Information.

Wetherill, G.B. (1981) *Intermediate statistical method,* London: Chapman and Hall.

Wheatcroft, P. (2005) 'Taxed at holes in the wall', *The Times,* 31 March, p. 41, London: Times Newspaper.

White, D. (2008) 'Financial: RBS to sell its stake in Tesco bank', *Daily Telegraph*, 2 June, (City pages), London: Telegraph Media Group.

Whiting, G.J. (1999) 'Monetize Your Bricks and Mortar Through a Sale-Leaseback', *Mergers & Acquisitions Journal,* 34, pp.9-13.

Whitmore, J. (2005) 'Leslau and Hunter's £100m care package', *Property Week*, 6 May, London: CMP Information. Available at: http://www.propertyweek.com/news/leslau-and-hunter's-£100m-care-package/3050601.article [Accessed 7 July 2011].

Wilding, M. (2010) 'Santander's £500m Mapeley buyback', *Property Week,* p.1, London: UBM Built Environment.

Williams, M. (2010) 'New proposals mean standards continue to confuse'. *Property Week,* p.34, London: UBM Built Environment.

Yeung, G. (2010) *Property & Facility Management Division: Chairman's Message*, p.34, Hong Kong: Hong Kong Institute of Surveyors. Available at: http://www.hkis.org.hk/hkis/cms/upload/Newsletter/nswl1239 [Accessed 7 March 2011].

Appendix I

Outline of the semi-structured interviews

1. The researcher introduced himself and then sought to establish credibility and rapport by giving the following information:

    - The researcher's professional qualifications
    - Nature of the researcher's professional practice and work
    - Details of the research project
    - Common acquaintances from professional practice

2. Respondents were encouraged to elaborate on their responses without prompting by the researcher, where possible, based on questions within the following interview guide:

    i Do you specialize in sale-and-leaseback by auction?

    ii Do you sell retail bank investment properties by auction?

    - If so, for whom?

    iii Could you outline those factors most likely to influence yield?

    - Where necessary, prompt to ask if the following have any influence:
        a) Bank company
        b) Region
        c) Lot size

    iv Do you think that any profile of building is more favoured?

3. The respondents were thanked for their valuable contribution and were told that they would be very welcome to have a copy of the thesis on completion.

Appendix II

The questionnaire pro forma

1. Default Section

This survey forms part of a post-graduate research project being conducted by Malvern Tipping at the Department of the Built Environment at the Anglia Ruskin University, Rivermead, Chelmsford. As a specialist in the field, you are one of a very small group invited to take part in this qualitative study. Your particpation would be very much appreciated.

The research is into sale-and-leaseback in the UK with the present focus on retail bank premises sold between 1997 and 2006. Therefore, the questions should be considered in the context of the period before the Credit Crunch and the effects that have resulted from it.

1. Your Name

[ ]

*** 2. Your professional qualifications**

[ ]

*** 3. Your position/job title**

[ ]

4. Name of your organization (and please give department)

[ ]

Please give your considered answers to the following, since as you are a specialist in the field, your responses will be an important part of the qualitative study to help frame the variables and categories to be used in the quantitative study.

*** 5. Please select from the list below the four factors that you consider to be most likely to have influenced investors' choice and/or yield when retail bank premises have been sold at auction on sale-and-leaseback terms. These choices should be made in the context of the market between 1997 and 2006, and therefore before the effects of the Credit Crunch.**

- ☐ Tenant bank company/brand
- ☐ Regional location
- ☐ Lot size (as in hammer price)
- ☐ Tenure
- ☐ Physical size of building
- ☐ Brand cross-over (i.e. two brands within a group in close proxmimity)
- ☐ Close proximity of branches of the same brand
- ☐ Age of the building
- ☐ Proximity of rival bank groups
- ☐ Other (Please specify in 6 below)

6. If you think that any factors other than those listed in Question 5 above have a greater bearing on investors' investment decisions when buying retail bank premises on sale-and-leaseback between 1997 and 2006, please list here.

[ ]

7. If you have any further comments or observations to make, these would be much appreciated.

[ ]

Appendix II (Continued)

The questionnaire pro forma (page 2)

Thank you very much for your particpation. Your input into this research is very much appreciated. Ideally, this survey should be completed on-line, in which case it will only take you three minutes. However, should you prefer to submit a hard copy, this should be sent to Malvern Tipping at Tipping's Barn, Uggeshall, Nr. Beccles, Suffolk, NR34 8BE. E-mail: admin@tipping-estates.co.uk

Appendix III

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
1	15 Jul 1997	Midland	London-M25	Ashford, Church Road	580000	45000	7.76	L	F	1997.03
2	15 Jul 1997	Midland	London-M25	Feltham, Hartington Rd W	206000	17500	8.50	S	F	1997.03
3	15 Jul 1997	Midland	London-M25	Hillingdon, Uxbridge Rd	305000	26500	8.69	M	F	1997.03
4	15 Jul 1997	Midland	London-M25	Sydenham, Sydenham Road	260000	18500	7.12	M	F	1997.03
5	15 Jul 1997	Midland	London-M25	Teddington, 42, High St	330000	25000	7.58	M	F	1997.03
6	15 Jul 1997	Midland	South-East	Ashtead, Surrey	198000	15000	7.58	S	F	1997.03
7	15 Jul 1997	Midland	South-East	Coulston, Surrey	205000	16000	7.80	S	F	1997.03
8	15 Jul 1997	Midland	South-East	Crowborough, East Sussex	148000	13250	8.95	S	F	1997.03
9	15 Jul 1997	Midland	South-East	Crownthorne, Berkshire	222000	18500	8.41	S	F	1997.03
10	15 Jul 1997	Midland	South-East	Edenbridge, Kent	191000	16000	8.38	S	F	1997.03
11	15 Jul 1997	Midland	South-East	Egham, Surrey	234000	18000	7.69	S	F	1997.03
12	15 Jul 1997	Midland	South-East	Fording bridge, Hants	158000	12500	7.91	S	F	1997.03
13	15 Jul 1997	Midland	South-East	Luton, Bedfordshire	162000	12500	7.72	S	F	1997.03
14	15 Jul 1997	Midland	South-East	Mitcham, Surrey	162000	12000	7.41	S	F	1997.03
15	15 Jul 1997	Midland	South-East	Pangbourne, Berkshire	235000	19000	8.09	S	F	1997.03
16	15 Jul 1997	Midland	South-East	Sandwich, Kent	159000	12500	7.86	S	F	1997.03
17	15 Jul 1997	Midland	South-East	Seaford, East Sussex	180000	14000	7.78	S	F	1997.03
18	15 Jul 1997	Midland	South-East	Shirley, Hampshire	232000	20500	8.84	S	F	1997.03
19	15 Jul 1997	Midland	South-East	Southsea, Hampshire	144000	12500	8.68	S	F	1997.03
20	15 Jul 1997	Midland	South-East	Storrington, West Sussex	186000	15000	8.06	S	F	1997.03
21	15 Jul 1997	Midland	South-East	Strood, Kent	620000	50000	8.06	L	F	1997.03
22	15 Jul 1997	Midland	South-East	Surbiton, Surrey	430000	30000	6.98	M	F	1997.03
23	15 Jul 1997	Midland	South-East	Tenterden, Kent	301000	24000	7.97	M	F	1997.03
24	15 Jul 1997	Midland	South-East	Tadley, Hampshire	212000	17000	8.02	S	F	1997.03
25	15 Jul 1997	Midland	South-East	Woolston, Hampshire	251000	22500	8.96	M	F	1997.03
26	15 Jul 1997	Midland	South-West	Amesbury, Wiltshire	240000	18500	7.71	S	F	1997.03
27	15 Jul 1997	Midland	South-West	Bedminster, Avon	276000	25000	9.06	M	F	1997.03

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
28	15 Jul 1997	Midland	South-West	Boscombe, Dorset	349000	25000	7.16	M	F	1997.03
29	15 Jul 1997	Midland	South-West	Bridgwater, Somerset	415000	32500	7.83	M	F	1997.03
30	15 Jul 1997	Midland	South-West	Brixham, Devon	258000	20000	7.75	M	F	1997.03
31	15 Jul 1997	Midland	South-West	Crediton, Devon	177000	15000	8.47	S	F	1997.03
32	15 Jul 1997	Midland	South-West	Cullumpton, Devon	132000	11000	8.33	S	F	1997.03
33	15 Jul 1997	Midland	South-West	Great Torrington, Devon	173500	13250	7.64	S	F	1997.03
34	15 Jul 1997	Midland	South-West	Holsworthy, Devon	138000	10000	7.25	S	F	1997.03
35	15 Jul 1997	Midland	South-West	Honiton, Devon	231000	18500	8.01	S	F	1997.03
36	15 Jul 1997	Midland	South-West	Lydney, Glos	181000	13500	7.46	S	F	1997.03
37	15 Jul 1997	Midland	South-West	Minehead, Somerset	335000	25000	7.46	M	F	1997.03
38	15 Jul 1997	Midland	South-West	Nailsworth, Glos	145000	11500	7.93	S	F	1997.03
39	15 Jul 1997	Midland	South-West	Okehampton, Devon	254000	20000	7.87	M	F	1997.03
40	15 Jul 1997	Midland	South-West	Shepton Mallet, Somerset	213500	18000	8.43	S	F	1997.03
41	15 Jul 1997	Midland	South-West	Totnes, Devon	210000	15000	7.14	S	F	1997.03
42	15 Jul 1997	Midland	South-West	Wellington, Somerset	276000	19500	7.07	M	F	1997.03
43	15 Jul 1997	Midland	South-West	Winscombe, Somerset	98000	7500	7.65	S	F	1997.03
44	15 Jul 1997	Midland	South-West	Winton, Dorset	373000	27500	7.37	M	F	1997.03
45	15 Jul 1997	Midland	East Anglia	Downham Market, Norfolk	200000	17000	8.50	S	F	1997.03
46	15 Jul 1997	Midland	East Anglia	Ipswich, Suffolk	136500	10000	7.33	S	F	1997.03
47	15 Jul 1997	Midland	West Mid	Coventry, Earlsdon St	192000	15000	7.81	S	F	1997.03
48	15 Jul 1997	Midland	West Mid	Newport, Shropshire	174000	14000	8.05	S	F	1997.03
49	15 Jul 1997	Midland	East Mid	Alfreton, Derbyshire	169000	15000	8.88	S	F	1997.03
50	15 Jul 1997	Midland	East Mid	Buxton, Derbyshire	208000	16500	7.93	S	F	1997.03
51	15 Jul 1997	Midland	East Mid	Cleethorpes, Lincolnshir	126000	11000	8.73	S	F	1997.03
52	15 Jul 1997	Midland	Wales	Aberaeron, Dyfed	107000	8000	7.48	S	F	1997.03
53	15 Jul 1997	Midland	Wales	Ammanford, Dyfed	117000	9000	7.69	S	F	1997.03
54	15 Jul 1997	Midland	Wales	Bala, Gwynedd	206000	17000	8.25	S	F	1997.03

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
55	15 Jul 1997	Midland	Wales	Caercwrle, Clwyd	76000	5500	7.24	S	F	1997.03
56	15 Jul 1997	Midland	Wales	Chirk, Clwyd	82000	7000	8.54	S	F	1997.03
57	15 Jul 1997	Midland	Wales	Fishguard, Dyfed	120000	10000	8.33	S	F	1997.03
58	15 Jul 1997	Midland	Wales	Flint, Clwyd	212000	17000	8.02	S	F	1997.03
59	15 Jul 1997	Midland	Wales	Llandeilo, Dyfed	80000	6000	7.50	S	F	1997.03
60	15 Jul 1997	Midland	Wales	Llandovery, Dyfed	115000	9000	7.83	S	F	1997.03
61	15 Jul 1997	Midland	Wales	Llanfair Caereinion	139000	11500	8.27	S	F	1997.03
62	15 Jul 1997	Midland	Wales	Llanfyllin, Powys	75000	5500	7.33	S	F	1997.03
63	15 Jul 1997	Midland	Wales	Llangollen, Clwyd	155500	11000	7.07	S	F	1997.03
64	15 Jul 1997	Midland	Wales	Llanidloes, Dyfed	94500	6500	6.88	S	F	1997.03
65	15 Jul 1997	Midland	Wales	Machynleth, Powys	109000	8000	7.34	S	F	1997.03
66	15 Jul 1997	Midland	Wales	Menai Bridge, Gwynedd	138000	12000	8.70	S	F	1997.03
67	15 Jul 1997	Midland	Wales	Monmouth, Gwent	180000	16000	8.89	S	F	1997.03
68	15 Jul 1997	Midland	Wales	Morriston, W.Glamorgan	161000	13500	8.39	S	F	1997.03
69	15 Jul 1997	Midland	Wales	Mumbles, West Glamorgan	231000	16000	6.93	S	F	1997.03
70	15 Jul 1997	Midland	Wales	Narberth, Dyfed	69000	5500	7.97	S	F	1997.03
71	15 Jul 1997	Midland	Wales	Pembroke, Dyfed	106000	8500	8.02	S	F	1997.03
72	15 Jul 1997	Midland	Wales	Porthcawl, Mid Glamorgan	205000	18000	8.78	S	F	1997.03
72	15 Jul 1997	Midland	Wales	Rhos-on-Sea, Clwyd	156000	13250	8.49	S	F	1997.03
74	15 Jul 1997	Midland	Wales	Risca, Gwent	88000	7000	7.95	S	F	1997.03
75	15 Jul 1997	Midland	Wales	Tywyn, Gwynedd	97000	8000	8.25	S	F	1997.03
76	15 Jul 1997	Midland	North-East	Acomb, York, North Yorks	120000	9000	7.50	S	F	1997.03
77	15 Jul 1997	Midland	North-East	Bradford, Dudley Hill	382000	28000	7.33	M	F	1997.03
78	15 Jul 1997	Midland	North-East	Brighouse, Wesr Yorkshir	172000	13500	7.85	S	F	1997.03
79	15 Jul 1997	Midland	North-East	Filey, North Yorkshire	210000	16000	7.62	S	F	1997.03
80	15 Jul 1997	Midland	North-East	Haltwhistle, Northumberl	86000	6500	7.56	S	F	1997.03
81	15 Jul 1997	Midland	North-East	Hemsworth, West Yorks	190000	16000	8.42	S	F	1997.03

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
82	15 Jul 1997	Midland	North-East	Hessle, East Yorkshire	180000	14500	8.06	S	F	1997.03
83	15 Jul 1997	Midland	North-East	Holmfirth, West Yorks	205000	17000	8.29	S	F	1997.03
84	15 Jul 1997	Midland	North-East	Hornsea, East Yorkshire	108000	9500	8.80	S	F	1997.03
85	15 Jul 1997	Midland	North-East	Ilkley, West Yorks	286000	22500	7.87	M	F	1997.03
86	15 Jul 1997	Midland	North-East	Morpeth, Northumberland	337000	28000	8.31	M	F	1997.03
87	15 Jul 1997	Midland	North-East	Pickering, North Yorks	165000	13500	8.18	S	F	1997.03
88	15 Jul 1997	Midland	North-East	Ponteland, Northumberlan	343500	27500	8.01	M	F	1997.03
89	15 Jul 1997	Midland	North-East	Richmond, North Yorks	411000	33500	8.15	M	F	1997.03
90	15 Jul 1997	Midland	North-East	Tadcaster, North Yorks	131000	10000	7.63	S	F	1997.03
91	15 Jul 1997	Midland	North-West	Aaspatria, Cumbria	87000	7000	8.05	S	F	1997.03
92	15 Jul 1997	Midland	North-West	Cheadle, Cheshire	193000	14500	7.51	S	F	1997.03
93	15 Jul 1997	Midland	North-West	Darwen, Lancashire	183000	13750	7.51	S	F	1997.03
94	15 Jul 1997	Midland	North-West	Kirkby Lonsdale, Cumbria	200000	14000	7.00	S	F	1997.03
95	15 Jul 1997	Midland	North-West	Poulton-le-Fylde, Lancs	276000	22500	8.15	M	F	1997.03
96	15 Jul 1997	Midland	North-West	Urmston, Gt.Manchester	261000	22500	8.62	M	F	1997.03
97	15 Jul 1997	Midland	North-West	Whitehaven, Cumbria	246000	20000	8.13	S	F	1997.03
98	15 Jul 1997	Midland	Wales	Queensferry, Clwyd		8000				1997.03
99	15 Jul 1997	Midland	East Mid	Uppingham, Leics		30000				1997.03
100	25 Feb 2004	BW	South-East	Hoddesdon Herts	396000	20000	5.50	M	F	2004.01
101	25 Feb 2004	BW	Wales	Carmarthen, Dyfed	365000	20000	5.50	M	F	2004.01
102	25 Feb 2004	BB	South-East	Camberley, Surrey	450000	24000	5.30	M	F	2004.01
103	25 Feb 2004	BB	South-East	Bicester, Oxfordshire	667000	35000	5.25	L	F	2004.01
104	25 Feb 2004	BB	North-East	Berverley, Yorks	715000	37500	5.24	L	F	2004.01
105	25 Feb 2004	BB	South-East	Winchster, Hampshire	1735000	95000	5.48	L	F	2004.01
106	25 Feb 2004	BW	South-West	Exeter, Devon	2875000	188500	6.56	L	F	2004.01
107	25 Feb 2004	BW	South-West	Knowle, Bristol	345000	20000	5.80	M	F	2004.01
108	25 Feb 2004	BB	Wales	Pontypridd, W. Glamorgan	570000	31000	5.44	L	F	2004.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
109	25 Feb 2004	BB	Wales	Swansea, Union Street	772000	45000	5.83	L	F	2004.01
110	25 Feb 2004	BW	Wales	Newport, Gwent	466000	25000	5.36	M	F	2004.01
111	25 Feb 2004	BB	South-West	Tavistock, Devon	344000	18000	5.20	M	F	2004.01
112	25 Feb 2004	BB	North-West	Lancaster	427000	20000	4.68	M	F	2004.01
113	25 Feb 2004	BW	South-East	Newport, Isle of Wight	495000	27500	5.55	M	F	2004.01
114	25 Feb 2004	BB	South-East	Bexleyheath, Kent	810000	46000	5.68	L	F	2004.01
115	25 Feb 2004	BW	South-West	Clevedon, Bristol	352000	17500	4.97	M	F	2004.01
116	25 Feb 2004	BW	South-West	Wellington, Somerset	402000	24000	5.97	M	L	2004.01
117	25 Feb 2004	BB	North-West	Whitehaven, Cumbria	780000	41000	5.26	L	F	2004.01
118	25 Feb 2004	BB	Scotland	Fraserburgh, Aberdeenshi	329000	19500	5.93	M	Feuhold	2004.01
119	25 Feb 2004	LTSB	Scotland	Saltcoats, Ayrshire		13000				2004.01
120	25 Feb 2004	BB	North-East	Stockton-on-Tees	640000	38000	5.94	L	L	2004.01
121	25 Feb 2004	BB	East Mid	Grimsby, Lincolnshire	725000	43000	5.93	L	F	2004.01
122	25 Feb 2004	BB	North-West	Bolton, Greater Manchest	1680000	88000	5.24	L	F	2004.01
123	25 Feb 2004	BB	North-West	Southport, Merseyside	1220000	55000	4.51	L	F	2004.01
124	25 Feb 2004	BB	South-East	Bromley, Kent	1530000	86500	5.65	L	F	2004.01
125	18 May 2004	HSBC	South-West	Brixham, Devon		24000			F	2004.02
126	18 May 2004	Barclay	East Anglia	St. Ives, Cambs		78500			F	2004.02
127	11 Oct 2004	BW	South-West	Liskeard, Cornwall	351000	19092	5.44	M	F	2004.04
128	11 Oct 2004	Derby B	West Mid	Coventry, Warks	631000	39000	6.19	L	F	2004.04
129	30 Nov 2004	BW	South-West	Shepton Mallet, Somerset	241000	12000	4.98	S	L	2004.04
130	30 Nov 2004	BW	South-West	Wellington, Somerset	415000	24000	5.78	M	L	2004.04
131	30 Nov 2004	LTSB	South-West	Boscombe, Bournemouth Do	610000	30000	4.92	L	F	2004.04
132	30 Nov 2004	YB	North-East	Rotherham, South Yorks	1616000	87500	5.41	L	F	2004.04
133	30 Nov 2004	LTSB	Scotland	Gorgie, Edinburgh	219000	14000	6.39	S	F	2004.04
134	01 Jul 2004	HSBC	South-East	Chessington, Surrey	372000	20000	5.38	M	F	2004.03
135	02 Mar 2005	BB	North-East	Pontefract, West Yorks	461000	22500	4.88	M	F	2005.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
136	02 Mar 2005	BB	East Mid	Derby, St.Peter's Street	1120000	50000	4.46	L	F	2005.01
137	02 Mar 2005	LTSB	Wales	Cardiff, Wellfield Road	930000	52250	5.62	L	L	2005.01
138	02 Mar 2005	Barclay	North-West	Wallasey, Wirral		72500			F	2005.01
139	02 Mar 2005	BB	North-West	Oldham, Gt. Manchester	1262500	70000	5.54	L	L	2005.01
140	02 Mar 2005	BB	North-East	Wakefield, West Yorks	1345000	73000	5.43	L	F	2005.01
141	02 Mar 2005	LTSB	North-West	Stockport, Gt Manchester	421000	27000	6.41	M	F	2005.01
142	02 Mar 2005	YB	North-West	Blackpool, Central Drive	356000	15000	4.21	M	F	2005.01
143	02 Mar 2005	Woolwic	Scotland	Kirkintilloch, Strathcly	435000	27800	6.39	M	F	2005.01
144	02 Mar 2005	BB	South-West	Swindon, Wiltshire	1236000	70000	5.66	L	F	2005.01
145	02 Mar 2005	BB	South-West	Exeter, Devon	1080000	53000	4.90	L	F	2005.01
146	02 Mar 2005	BB	South-East	Guildford, Surrey	2630000	112500	4.28	L	F	2005.01
147	23 May 2005	Abbey	London-M25	Dagenham, Essex	423000	24000	5.67	M	F	2005.02
148	23 May 2005	HFX	South-West	Gloucester, Glos	1800000	95000	5.28	L	F	2005.02
149	19 Apr 2005	BW	South-West	Salisbury, Wiltshire	2210000	110000	4.98	L	F	2005.02
150	05 Jul 2005	RBS	West Mid	Wolverhampton, Princes S	959000	44250	4.61	L	F	2005.03
151	05 Jul 2005	A&L	Scotland	Dumfries	642000	37150	5.79	L	Heritabl	2005.03
152	05 Jul 2005	BB	North-West	Oldham, Gt Manchester	1279000	70000	5.47	L	L	2005.03
153	05 Jul 2005	NatWest	North-West	St Annes-on-Sea, Lancs	1064000	69839	5.65	L	F	2005.03
154	06 Oct 2005	Abbey	South-West	Boscombe, Bournemouth, D		41500			F	2005.04
155	06 Oct 2005	LTSB	South-East	Hastings, Silverhill, Su	430000	22000	5.12	M	F	2005.04
156	06 Oct 2005	LTSB	West Mid	Wednesbury	256000	10000	3.91	M	F	2005.04
157	30 Nov 2005	RBS	South-East	Fareham, Hampshire	433000	25000	5.77	M	F	2005.04
158	30 Nov 2005	LTSB	South-West	Liskeard, Cornwall	475000	26000	5.47	M	F	2005.04
159	30 Nov 2005	HSBC	South-West	Bedminster, Avon	454000	26000	5.73	M	F	2005.04
160	30 Nov 2005	HSBC	London-M25	East Ham, London	1075000	56000	5.21	L	F	2005.04
161	30 Nov 2005	LTSB	Scotland	Dundee, Albert Street	500000	30000	6.00	M	Heritabl	2005.04
162	30 Nov 2005	Clyde	East Mid	Bedford, Bedfordshire	1100000	57500	5.23	L	F	2005.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
163	30 Nov 2005	HFX	West Mid	Kidderminster, Worcs	1733000	85000	4.90	L	F	2005.04
164	30 Nov 2005	HSBC	Wales	Bala, Gwynedd	332000	17000	5.12	M	F	2005.04
165	30 Nov 2005	Barclay	Wales	Porth, Rhondda	276000	15000	5.43	M	F	2005.04
166	09 Feb 2006	Barclay	South-East	Crowborough, E.Sussex	677000	30500	4.51	L	F	2006.01
167	09 Feb 2006	LTSB	South-West	Swindon, Gorse Hill	478000	25000	5.23	M	F	2006.01
168	09 Feb 2006	Woolwic	South-East	Carshalton WL	550000	27500	5.00	L	F	2006.01
169	28 Mar 2006	Barclay	South-East	Chipping Norton, Oxon	780000	36900	4.73	L	F	2006.01
170	28 Mar 2006	Barclay	South-East	Princes Risborough, Buck	515000	23650	4.59	L	F	2006.01
171	28 Mar 2006	Barclay	South-East	Crowthorne, Berkshire	735000	38200	5.20	L	F	2006.01
172	28 Mar 2006	BB	North-East	Harrogate, Yorks	1260000	90000	7.14	L	L	2006.01
173	28 Mar 2006	BB	West Mid	Burton-on-Trent, Staffs	665000	32000	4.81	L	F	2006.01
174	28 Mar 2006	BB	East Mid	Oadby, Leicestershire	366000	15000	4.10	M	F	2006.01
175	28 Mar 2006	Barclay	West Mid	Newport, Shropshire		21963			F	2006.01
176	28 Mar 2006	BB	North-West	Chorley, Lancashire	571000	23000	4.03	L	F	2006.01
177	28 Mar 2006	BB	South-East	Hove, East Sussex	558000	21000	3.76	L	F	2006.01
178	28 Mar 2006	Barclay	South-East	Tilehurst, Berkshire	735000	32200	4.38	L	F	2006.01
179	28 Mar 2006	Barclay	South-East	Leigh-on-Sea, Essex	505000	20000	3.96	L	F	2006.01
180	28 Mar 2006	BB	East Mid	Worksop, Nottinghamshire	591000	30500	5.16	L	F	2006.01
181	28 Mar 2006	BB	Scotland	Glasgow, Byres Road	2110000	87500	4.15	L	Heritabl	2006.01
182	28 Mar 2006	Barclay	North-West	Middleton, Manchester	862000	40000	4.64	L	F	2006.01
183	28 Mar 2006	Barclay	East Mid	Brackley. Northants	670000	34500	5.15	L	F	2006.01
184	28 Mar 2006	Barclay	South-East	Tadley, Hampshire	610000	29500	4.84	L	F	2006.01
185	28 Mar 2006	Barclay	South-East	Potters Bar, Herts	376000	17500	4.65	M	F	2006.01
186	28 Mar 2006	Barclay	South-East	Billinghurst, W. Sussex	705000	30600	4.34	L	F	2006.01
187	28 Mar 2006	Barclay	South-East	Bicester, Oxfordshire	1010000	55000	5.45	L	F	2006.01
188	28 Mar 2006	Barclay	South-East	Longfield, Kent	463000	22000	4.75	M	F	2006.01
189	28 Mar 2006	Barclay	East Mid	Sleaford, Lincolnshire	685000	31200	4.55	L	F	2006.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
190	28 Mar 2006	Barclay	North-East	Pontefract, W. Yorks	1355000	72500	5.35	L	F	2006.01
191	28 Mar 2006	Barclay	South-West	Plympton, Devon	460000	20000	4.35	M	F	2006.01
192	28 Mar 2006	Barclay	South-East	Deal, Kent	480000	24000	5.00	M	F	2006.01
193	28 Mar 2006	Barclay	North-East	Dewsbury, West Yorks	1185000	62500	5.27	L	F	2006.01
194	28 Mar 2006	Barclay	North-East	Shipley, West Yorkshire	502500	22660	4.51	L	F	2006.01
195	28 Mar 2006	Barclay	Wales	Neath, West Glamorgan	530000	23000	4.34	L	F	2006.01
196	25 May 2006	Barclay	South-East	Esher, Surrey	1475000	57500	3.90	L	F	2006.02
197	25 May 2006	Barclay	London-M25	Bow, Roman Road	710000	27000	3.80	L	F	2006.02
198	25 May 2006	Barclay	London-M25	Upminster, Essex	740000	32600	4.41	L	F	2006.02
199	25 May 2006	Barclay	London-M25	Feltham, Middlesex	1280000	59900	4.68	L	F	2006.02
200	25 May 2006	Barclay	South-West	Broadstone, Dorset		43000			F	2006.02
201	25 May 2006	LTSB	North-West	Nelson, Lancashire	462500	27000	5.84	M	F	2006.02
202	25 May 2006	Barclay	North-East	Byker, Newcastle-u-Tyne	688000	35000	5.09	L	F	2006.02
203	25 May 2006	Barclay	South-West	Sidmouth, Devon		30000			F	2006.02
204	25 May 2006	Barclay	South-West	Honiton, Devon	801000	35000	4.37	L	F	2006.02
205	25 May 2006	Britann	South-East	Worthing, West Sussex	625000	33000	5.28	L	L	2006.02
206	25 May 2006	Barclay	East Anglia	Diss, Norfolk	765000	32500	4.25	L	F	2006.02
207	25 May 2006	Barclay	East Anglia	Thetford, Norfolk	876000	45000	5.14	L	F	2006.02
208	25 May 2006	Barclay	South-East	Farnborough, Kent	526000	25500	4.85	L	F	2006.02
209	25 May 2006	Woolwic	London-M25	Grays, Essex	780000	38000	4.87	L	F	2006.02
210	25 May 2006	Barclay	South-East	Yateley, Hampshire	696000	36000	5.17	L	F	2006.02
211	03 Jul 2006	HSBC	South-East	Hoddesdon, Herts	424000	20000	4.72	M	F	2006.03
212	03 Jul 2006	HSBC	South-East	Gerrards Cross, Bucks		37500			F	2006.03
213	03 Jul 2006	Barclay	London-M25	Ponders End, Middlesex	1210000	57200	4.73	L	F	2006.03
214	03 Jul 2006	Barclay	South-West	Truro, Cornwall	1250000	61400	4.91	L	F	2006.03
215	03 Jul 2006	HSBC	South-East	Lymington. Hampshire	1454000	60000	4.13	L	F	2006.03
216	03 Jul 2006	HSBC	South-East	Sidcup, Kent	754000	40000	5.31	L	F	2006.03

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
217	03 Jul 2006	Barclay	North-East	Hexham, Northumberland	751000	36300	4.83	L	F	2006.03
218	03 Jul 2006	HSBC	West Mid	Leek, Staffordshire	550000	28500	5.18	L	F	2006.03
219	03 Jul 2006	HSBC	North-West	Chorley, Lancashire	605000	30000	4.96	L	F	2006.03
220	03 Jul 2006	HSBC	North-West	Altricham, Cheshire	1410000	75000	5.32	L	F	2006.03
221	03 Jul 2006	Other	North-East	Rotherham, Effingham St		25250			F	2006.03
222	03 Jul 2006	Barclay	London-M25	Wembley, Middlesex	1171000	40000	3.42	L	F	2006.03
223	03 Jul 2006	HSBC	South-East	Thame, Oxfordshire	440000	22500	5.11	M	F	2006.03
224	03 Jul 2006	HSBC	South-West	Malmesbury, Wiltshire	605000	30000	4.96	L	F	2006.03
225	03 Jul 2006	HSBC	West Mid	Newcastle-Under-Lyme	1191000	70000	5.88	L	F	2006.03
226	03 Jul 2006	HSBC	North-West	Blackpool, Oxford Square	630000	32500	5.16	L	F	2006.03
227	03 Jul 2006	HSBC	North-East	Armley, West Yorkshire	385000	20000	5.19	M	F	2006.03
228	03 Jul 2006	HSBC	West Mid	Tamworth, Staffordshire	725000	37500	5.17	L	F	2006.03
229	03 Jul 2006	HSBC	East Mid	Belper, Derbyshire	461000	24000	5.21	M	F	2006.03
230	03 Jul 2006	HSBC	Wales	Cardiff, Albany Road	740000	37500	5.07	L	F	2006.03
231	03 Jul 2006	HSBC	Wales	Mold, Flintshire	1053000	61000	5.79	L	F	2006.03
232	08 Jul 2004	BW	South-West	Henleaze, Bristol	400000	22600	5.65	M	F	2004.03
233	08 Jul 2004	HSBC	South-East	Reading, Berkshire		72500			F	2004.03
234	08 Jul 2004	Other	South-West	Exeter, Devon		25000			L	2004.03
235	08 Jul 2004	LTSB	South-East	Chelmsford, Essex	2060000	102500	4.98	L	F	2004.03
236	08 Jul 2004	LTSB	South-East	Gillingham, Kent	378000	20000	5.29	M	F	2004.03
237	25 Mar 2003	Abbey	Wales	Caernarfon	161000	11250	6.99	S	F	2003.01
238	25 Mar 2003	Barclay	South-East	Witham, Essex		43500			F	2003.01
239	25 Mar 2003	BW	South-West	Blandford Forum, Dorset	175000	10000	5.71	S	F	2003.01
240	25 Mar 2003	LTSB	East Mid	Grantham, Lincolnshire		50000			F	2003.01
241	25 Mar 2003	LTSB	South-West	Wareham, Dorset	540000	35000	6.48	L	F	2003.01
242	03 Dec 2003	LTSB	Scotland	Kilmarnock, Strathclyde	1150000	80000	6.96	L	Feuhold	2003.04
243	03 Dec 2003	LTSB	South-East	Bletchley, Milton Keynes	675000	42500	6.30	L	F	2003.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
244	03 Dec 2003	LTSB	South-West	Swindon, Gorse Hill	402000	25000	6.22	M	F	2003.04
245	03 Dec 2003	LTSB	West Mid	Netherton, Dudley, WM	125000	7500	6.00	S	F	2003.04
246	03 Dec 2003	LTSB	South-East	Redhill, Surrey		114500			F	2003.04
247	03 Dec 2003	LTSB	West Mid	Longbridge, West Midland	204000	10000	4.90	S	F	2003.04
248	08 Feb 2005	BB	South-West	Tavistock, Devon	285000	18600	6.53	M	F	2005.01
249	06 Jul 2005	LTSB	South-East	Princes Risborough, Buck	615000	30000	4.88	L	F	2005.03
250	06 Jul 2005	LTSB	South-East	Edenbridge, Kent	415000	20000	4.82	M	F	2005.03
251	06 Jul 2005	LTSB	Wales	Llandrindod Wells	282500	12500	4.42	M	F	2005.03
252	06 Jul 2005	LTSB	South-East	Haslemere, Shottermill	265000	13500	5.09	M	F	2005.03
253	06 Jul 2005	LTSB	South-East	Shoreham, West Sussex	342000	14000	4.09	M	F	2005.03
254	06 Jul 2005	LTSB	South-East	Silverhill, Hastings	365000	22000	6.03	M	F	2005.03
255	06 Jul 2005	LTSB	South-West	Horfield, Bristol	600000	19000	3.17	L	F	2005.03
256	06 Jul 2005	LTSB	North-West	Heswall, Merseyside	605000	30000	4.96	L	F	2005.03
257	06 Jul 2005	LTSB	South-East	New Milton, Hants	437500	20000	4.57	M	F	2005.03
258	06 Jul 2005	LTSB	South-East	Petts Wood, Kent	400000	17750	4.44	M	F	2005.03
259	06 Jul 2005	LTSB	South-East	Leagrave, Luton, Beds	395000	18650	4.72	M	F	2005.03
260	06 Jul 2005	LTSB	South-East	Bookham, Surrey	555000	26700	4.81	L	F	2005.03
261	06 Jul 2005	LTSB	South-East	Dartford, Kent	917500	49000	5.34	L	F	2005.03
262	06 Jul 2005	LTSB	South-East	Gosport, Hampshire	1035000	60000	5.80	L	F	2005.03
263	06 Jul 2005	LTSB	South-West	Cullumpton, Devon	325000	17000	5.23	M	F	2005.03
264	06 Jul 2005	LTSB	East Mid	Rushden, Northants	407500	21500	5.28	M	F	2005.03
265	06 Jul 2005	LTSB	South-East	Maidstone LR, Kent	370000	18000	4.86	M	F	2005.03
266	06 Jul 2005	LTSB	West Mid	Sparkhill	480000	18000	3.75	M	F	2005.03
267	06 Jul 2005	LTSB	North-East	Gateshead	587500	19000	3.23	L	F	2005.03
268	06 Jul 2005	LTSB	West Mid	Wellington, Shropshire	500000	28000	5.60	M	F	2005.03
269	06 Jul 2005	LTSB	London-M25	Southall, Middlesex	3050000	100000	3.28	L	L	2005.03
270	22 Mar 2005	LTSB	North-West	Barrow-in-Furness, Cumbr		77500			F	2005.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
271	08 Feb 2005	LTSB	South-East	Liphook, Hampshire	650000	37000	5.69	L	F	2005.01
272	08 Feb 2005	LTSB	East Anglia	Wisbech, Cambs	430000	19000	4.42	M	F	2005.01
273	08 Feb 2005	LTSB	North-East	Rothbury, Northumberland	340000	11000	3.24	M	F	2005.01
274	13 Oct 2004	LTSB	South-West	Brockworth, Glos	292500	16000	5.47	M	F	2004.04
275	19 May 2004	LTSB	East Mid	Leicester SAR	465000	22500	4.84	M	F	2004.02
276	19 May 2004	LTSB	South-East	Coulsdon, Surrey	545000	28000	5.14	L	F	2004.02
277	19 May 2004	LTSB	East Mid	Melton Mowbray, Leics	732500	42500	5.80	L	F	2004.02
278	25 Mar 2004	LTSB	West Mid	Lichfield, Staffordshire	950000	65250	6.87	L	F	2004.01
279	25 Mar 2004	LTSB	South-West	Totnes, Devon	1020000	49500	4.85	L	F	2004.01
280	25 Mar 2004	LTSB	South-East	Maidenhead, Berks	1900000	200000	6.45	L	F	2004.01
281	25 Mar 2004	LTSB	South-East	Maidstone, Kent	3410000	200000	5.87	L	F	2004.01
282	25 Mar 2004	LTSB	West Mid	Walsall, West Midlands	2700000	180000	6.67	L	F	2004.01
283	25 Mar 2004	LTSB	North-East	Halifax, Yorks	1050000	62000	5.90	L	F	2004.01
284	25 Mar 2004	LTSB	South-East	Hadleigh, Essex	400000	18500	4.63	M	F	2004.01
285	25 Mar 2004	LTSB	East Mid	Bulwell, Nottinghamshire	650000	42000	6.46	L	F	2004.01
286	25 Mar 2004	LTSB	North-West	Moston, Manchester	330000	22000	6.67	M	F	2004.01
287	25 Mar 2004	LTSB	South-West	Crownhill, Plymouth	181000	9500	5.25	S	F	2004.01
288	10 Feb 2004	LTSB	South-East	Hythe, Hampshire	495000	22500	4.55	M	F	2004.01
289	10 Feb 2004	LTSB	South-West	Winton, Bournemouth	600000	35000	5.83	L	F	2004.01
290	10 Feb 2004	LTSB	North-West	GCS Liverpool	1500000	110000	7.33	L	F	2004.01
291	10 Feb 2004	LTSB	West Mid	Evesham, Worcs	1115000	60000	5.38	L	F	2004.01
292	10 Feb 2004	LTSB	Scotland	Dundee, Albert Street		56000			Feuhold	2004.01
293	03 Dec 2003	LTSB	South-East	Wingham, Kent	176000	7500	4.26	S	F	2003.04
294	03 Dec 2003	LTSB	West Mid	Stratford-upon-Avon	2450000	109000	4.45	L	F	2003.04
295	03 Dec 2003	LTSB	East Mid	Grantham, Lincolnshire	650000	50000	7.69	L	F	2003.04
296	21 May 2003	YB	North-West	Stockport	1530000	82000	5.36	L	F	2003.02
297	25 Mar 2003	Clyde	Scotland	Larbert	250000	13500	5.40	S	Heritabl	2003.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
298	06 Jul 2005	Clyde	Scotland	Largs, North Ayrshire	585000	34400	5.88	L	Heritabl	2005.03
299	21 May 2003	NatWest	South-East	Hitchen, Hertfordshire	470000	20825	4.43	M	F	2003.02
300	03 dec 2003	NatWest	London-M25	Dagenham, Essex		20000			F	2003.04
301	25 Mar 2003	Barclay	South-East	Witham, Essex		43500			F	2003.01
302	10 Feb 2004	Barclay	South-East	Dorking, Surrey	1710000	94000	5.50	L	F	2004.01
303	10 Feb 2004	Barclay	South-East	South Lancing	620000	37900	6.11	L	F	2004.01
304	19 May 2004	Barclay	South-East	Stopsley, Beds	500000	28800	5.76	M	F	2004.02
305	19 May 2004	Barclay	North-West	Wallasey, Merseyside		72500			F	2004.02
306	19 May 2004	Barclay	South-East	Purley, Surrey	580000	30000	5.17	L	F	2004.02
307	10 Feb 2004	HSBC	South-East	Chatham, Kent	1000000	70000	7.00	L	F	2004.01
308	13 Oct 2004	HSBC	South-West	Cullumpton, Devon	224000	11500	5.13	S	F	2004.04
109	19 May 2005	HSBC	South-West	Boscombe, Bournemouth	470000	25000	5.32	M	F	2005.02
310	13 Oct 2004	Barclay	South-East	Chatham, Kent	2350000	150000	6.38	L	F	2004.04
311	13 Oct 2004	Barclay	South-West	Gloucester, Glos	1060000	53000	5.00	L	F	2004.04
312	08 Feb 2004	Barclay	London-M25	Wandsworth High Street		60000			L	2004.01
313	22 Mar 2005	BW	South-West	Gloucester, Glos	1500000	88372	5.89	L	F	2005.01
314	06 Jul 2005	BW	South-West	Chippenham, Wilts	725000	42000	5.79	L	F	2005.03
315	10 Oct 2005	Barclay	East Anglia	Long Stratton, Norfolk	310000	18450	5.95	M	F	2005.04
316	10 Oct 2005	Barclay	East Anglia	Framlingham, Suffolk	432000	20300	4.70	M	F	2005.04
317	10 Oct 2005	RBS	South-East	Ashford, Kent 1	1100000	64000	5.82	L	F	2005.04
318	10 Oct 2005	LTSB	West Mid	Stoke-on-Trent SS	242500	10000	4.12	S	F	2005.04
319	10 Oct 2005	Barclay	South-East	West Mersea, Essex	285000	14850	5.21	M	F	2005.04
320	04 Nov 2004	Barclay	London-M25	Grays, Essex	235000	14000	5.98	S	F	2004.04
321	26 Sep 2004	LTSB	Wales	Pontypool, Gwent	241000	10000	4.15	S	F	2004.03
322	14 Jul 2004	LTSB	North-East	Sunderland, Tyneside	630000	40000	6.35	L	F	2004.03
323	02 Jun 2004	COOP	East Mid	Cheapside, Leicester		13945			F	2004.02
324	20 Apr 2004	RBS	North-West	Burnley, Lancashire	690000	38500	5.58	L	F	2004.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
325	20 Apr 2004	LTSB	East Anglia	Clare, Suffolk	141000	7000	4.96	S	F	2004.02
326	20 Apr 2004	LTSB	North-West	New Ferry, Wirral	250000	12600	5.04	S	F	2004.02
327	20 Feb 2003	LTSB	North-West	Bolton, Lancashire	890000	68000	7.64	L	F	2003.01
328	19 Mar 2002	LTSB	Wales	Brecon, Powys	295000	21002	7.12	M	F	2002.01
329	20 Feb 2003	LTSB	North-East	Ryton, Newcastle-u-Tyne		5000			F	2003.01
330	02 May 2002	BW	South-West	Glastonbury, Somerset	370000	22920	6.19	M	F	2002.02
331	02 May 2002	BW	South-East	Maidenhead, Berkshire	607000	43128	7.11	L	F	2002.02
332	02 May 2002	HSBC	South-West	Kingsbridge, Devon	276000	15500	5.62	M	F	2002.02
333	02 May 2002	HSBC	South-West	Okehampton, Devon	350000	20000	5.71	M	F	2002.02
334	02 May 2002	LTSB	South-West	Barnstable, Devon	765000	60000	7.84	L	F	2002.02
335	02 May 2002	LTSB	South-East	Bexhill-on-Sea, E.Sussex	800000	57500	7.19	L	F	2002.02
336	02 May 2002	LTSB	Scotland	Gorgie, Edinburgh, M.Lot	181000	14000	7.73	S	Feuhold	2002.02
337	03 Apr 2003	HSBC	North-East	Richmond, North Yorkshir	656000	40000	6.10	L	F	2003.02
338	03 Apr 2003	LTSB	West Mid	Stourbridge, Lye, WM	164000	10000	6.10	S	F	2003.02
339	03 Apr 2003	HFX	Scotland	Dumbarton, Strathclyde	1220000	75000	6.15	L	Feuhold	2003.02
340	07 Jul 2003	HSBC	South-East	Faversham, Kent	314000	18500	5.89	M	F	2003.03
341	07 Jul 2003	LTSB	East Mid	Heanor, Derbyshire	315000	18000	5.71	M	F	2003.03
342	07 Jul 2003	LTSB	Wales	Shotton, Clwyd	420000	29000	6.90	M	F	2003.03
343	07 Oct 2003	Abbey	North-East	Stockton-on-Tees	907000	55000	6.06	L	F	2003.04
344	07 Oct 2003	BB	North-West	Blackburn, Lancashire	835000	41000	4.91	L	F	2003.04
345	07 Oct 2003	BB	South-East	Brighton, East Sussex	856000	45000	5.26	L	F	2003.04
346	07 Oct 2003	BB	North-West	Burnley, Lancashire	563000	30000	5.33	L	L	2003.04
347	07 Oct 2003	BB	South-East	Chatham, Kent	865000	49000	5.66	L	F	2003.04
348	07 Oct 2003	BB	South-East	Chelmsford, Essex	1210000	63000	5.21	L	F	2003.04
349	07 Oct 2003	BB	West Mid	Evesham, Worcs	581000	35000	6.02	L	F	2003.04
350	07 Oct 2003	BB	North-East	Hull, Paragon Hill	813000	45000	5.54	L	F	2003.04
351	07 Oct 2003	BB	South-East	Sutton, Surrey	1000000	51000	5.10	L	F	2003.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
352	07 Oct 2003	BB	South-East	Wokingham, Berks	580000	30000	5.17	L	F	2003.04
353	07 Oct 2003	BB	South-East	Worthing, West Sussex	568000	25000	4.40	L	F	2003.04
354	07 Oct 2003	NatWest	London-M25	Bethnal Green, London Rd	491000	37000	7.54	M	L	2003.04
355	09 Jul 2002	A&L	East Anglia	Norwich, Norfolk	1102000	72000	6.53	L	F	2002.03
356	09 Jul 2002	LTSB	East Mid	Arnold, Nottingham	600000	43000	7.17	L	F	2002.03
357	09 Jul 2002	LTSB	Wales	Cardiff, Clifton Street	245000	17500	7.14	S	F	2002.03
358	09 Jul 2002	LTSB	London-M25	Ruislip, Middlesex	704000	43500	6.18	L	F	2002.03
359	09 Jul 2002	LTSB	Wales	Swansea, Craddock St		84342			L	2002.03
360	09 Jul 2002	NatWest	North-East	Firth Park, Sheffield		17300			F	2002.03
361	12 Jul 2001	Barclay	London-M25	South Harrow, Middlesex		192800			F	2001.03
362	07 Oct 2003	BB	South-East	Horsham, Carfax, W.Susse	561000	31500	5.61	L	F	2003.04
363	16 Oct 2001	Abbey	North-East	Leeds, Park Row, N.Yorks	1250000	107500	8.60	L	L	2001.04
364	16 Oct 2001	LTSB	East Anglia	Newmarket, Suffolk	730000	52500	7.19	L	F	2001.04
365	18 May 2004	Barclay	East Anglia	St. Ives, Cambs		78500			F	2004.02
366	18 May 2004	HSBC	South-East	Ashford, Middlesex, Chur		51000			F	2004.02
367	21 Nov 2001	LTSB	South-East	Abbotts Langley, Herts	275000	13500	4.91	M	F	2001.04
368	21 Nov 2001	LTSB	West Mid	Alcester, Warwickshire	280000	15000	5.36	M	F	2001.04
369	21 Nov 2001	LTSB	South-West	Ashburton, Devon	170000	8000	4.71	S	F	2001.04
370	21 Nov 2001	LTSB	North-West	Ashton, Preston, Lancs	183000	13500	7.38	S	F	2001.04
371	21 Nov 2001	LTSB	West Mid	Aston, Long Acre, WM	250000	14000	5.60	S	F	2001.04
372	21 Nov 2001	LTSB	South-East	Barton-le-Clay	202000	13000	6.44	S	F	2001.04
373	21 Nov 2001	LTSB	North-West	Bebington, Wirral	115000	7500	6.52	S	F	2001.04
374	21 Nov 2001	LTSB	North-East	Birtley, Tyne & Wear	140000	9000	6.43	S	F	2001.04
375	21 Nov 2001	LTSB	South-West	Bovey Tracey, Devon	171000	9500	5.56	S	F	2001.04
376	21 Nov 2001	LTSB	West Mid	Bromyard, Herefordshire	241000	16000	6.64	S	F	2001.04
377	21 Nov 2001	LTSB	Wales	Brynmawr, Gwent	111000	6500	5.86	S	F	2001.04
378	21 Nov 2001	LTSB	Wales	Canton, Cardiff	276000	20000	7.25	M	F	2001.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
379	21 Nov 2001	LTSB	South-West	Chipping Sodbury, Glos	276000	15000	5.62	M	F	2001.04
380	21 Nov 2001	LTSB	North-East	Cleckheaton, West Yorks	165000	11000	6.67	S	F	2001.04
381	21 Nov 2001	LTSB	West Mid	Donnington, Telford	86000	5000	5.81	S	F	2001.04
382	21 Nov 2001	LTSB	South-East	East Horsley, Surrey	261000	12500	4.79	M	F	2001.04
383	21 Nov 2001	LTSB	London-M25	Finsbury Park, Seven Sis	371000	17500	4.72	M	F	2001.04
384	21 Nov 2001	LTSB	South-East	Fordingbridge, Hants	250000	14000	5.60	S	F	2001.04
385	21 Nov 2001	LTSB	South-East	Freshwater, Isle of Wigh	163000	8500	5.21	S	F	2001.04
386	21 Nov 2001	LTSB	South-West	Hanham, Bristol	192000	13500	7.03	S	F	2001.04
387	21 Nov 2001	LTSB	South-East	Hassocks, West Sussex	194000	12500	6.44	S	F	2001.04
388	21 Nov 2001	LTSB	North-East	Heaton, Newcastle-upon-T	378000	28250	7.47	M	F	2001.04
389	21 Nov 2001	LTSB	South-West	Highworth, Wiltshire	235000	11000	4.68	S	F	2001.04
390	21 Nov 2001	LTSB	South-West	Holsworthy, Devon	190000	10000	5.26	S	F	2001.04
391	21 Nov 2001	LTSB	East Mid	Horncastle, Lincolnshire	210000	14000	6.67	S	F	2001.04
392	21 Nov 2001	LTSB	South-West	Hugh Town, Scilly Isles	176000	10100	5.74	S	F	2001.04
393	21 Nov 2001	LTSB	West Mid	Stourbridge, Lye, WM	150000	10000	6.67	S	F	2001.04
394	21 Nov 2001	LTSB	West Mid	Market Drayton, Shrops	255000	15000	5.88	M	F	2001.04
395	21 Nov 2001	LTSB	Wales	Mountain Ash, M.Galmorga	202000	13750	6.81	S	F	2001.04
396	21 Nov 2001	LTSB	Wales	Mumbles, W.Glamorgan	315000	17500	5.56	M	F	2001.04
397	21 Nov 2001	LTSB	South-West	Nailsworth, Glos	256000	13500	5.27	M	F	2001.04
398	21 Nov 2001	LTSB	West Mid	Netherton, Dudley	110000	7500	6.82	S	F	2001.04
399	21 Nov 2001	LTSB	South-East	Odiham, Hampshire	403000	20000	4.96	M	F	2001.04
400	21 Nov 2001	LTSB	South-West	Oldfield Park, Bath	200000	10250	5.13	S	F	2001.04
401	21 Nov 2001	LTSB	Wales	Port Talbot, W.Glamorgan	504000	35000	6.94	L	F	2001.04
402	21 Nov 2001	LTSB	South-West	Plympton, Devon	346000	20000	5.78	M	F	2001.04
405	21 Nov 2001	LTSB	North-East	Ryton, Newcastle-upon-Ty	80000	5000	6.25	S	F	2001.04
406	21 Nov 2001	LTSB	South-West	South Molton, Devon	221000	12500	5.66	S	F	2001.04
407	21 Nov 2001	LTSB	South-West	Saltash, Cornwall	226000	15000	6.64	S	F	2001.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
408	21 Nov 2001	LTSB	East Anglia	Great Shelford, Cambs	270000	13500	5.00	M	F	2001.04
407	21 Nov 2001	LTSB	West Mid	Shipston-on-Stour, Warks	313000	16500	5.27	M	F	2001.04
408	21 Nov 2001	LTSB	East Mid	Splisby, Lincolnshire	145000	9000	6.21	S	F	2001.04
409	21 Nov 2001	LTSB	South-West	St Austell MCR, Cornwall	142000	7500	5.28	S	F	2001.04
410	21 Nov 2001	LTSB	South-West	St Blazey, Cornwall	164000	8750	5.34	S	F	2001.04
411	21 Nov 2001	LTSB	South-West	Stonehouse, Glos	300000	20000	6.67	M	F	2001.04
412	21 Nov 2001	LTSB	West Mid	Studley, Warwickshire	183000	13500	7.38	S	F	2001.04
413	21 Nov 2001	LTSB	South-West	Swanage, Dorset	287000	20000	6.97	M	F	2001.04
414	21 Nov 2001	LTSB	Wales	Swansea, Craddock Street	1030000	85000	8.25	L	L	2001.04
415	21 Nov 2001	LTSB	Scotland	Templehall, Kirkcaldy	187000	13000	6.95	S	Feuhold	2001.04
416	21 Nov 2001	LTSB	West Mid	Tutbury, Staffordshire	176000	11500	6.53	S	F	2001.04
417	21 Nov 2001	LTSB	South-West	Westbury, Wiltshire	229000	13500	5.90	S	F	2001.04
418	21 Nov 2001	LTSB	South-West	Williton, Somerset	138000	7500	5.43	S	F	2001.04
419	21 Nov 2001	LTSB	South-West	Wilton, Salisbury, Wilts	250000	15000	6.00	S	F	2001.04
420	21 Nov 2001	LTSB	South-East	Worthing GVA, W. Sussex	262000	15000	5.73	M	F	2001.04
421	21 Nov 2001	LTSB	South-West	Yatton, Somerset	117000	7500	6.41	S	F	2001.04
422	25 Feb 2002	BW	South-West	Teignmouth, Devon	485000	34128	7.04	M	F	2002.01
423	25 Feb 2002	BW	South-West	Weston-super-Mare	765000	56780	7.42	L	F	2002.01
424	25 Feb 2002	Clyde	Scotland	Aberdeen, Ellon Rd, Gram	203000	17750	8.75	S	Feuhold	2002.01
425	25 Feb 2002	Clyde	Scotland	Aberdeen GNR, Grampian	240000	24800	10.33	S	Feuhold	2002.01
426	25 Feb 2002	Clyde	Scotland	Dingwall, Highland	306500	24240	7.91	M	Feuhold	2002.01
427	25 Feb 2002	NatWest	East Anglia	King's Lynn, Hardwick Rd	610000	42225	6.92	L	F	2002.01
428	25 Feb 2002	Woolwic	West Mid	Stafford, Staffs	1100000	73500	6.68	L	F	2002.01
429	25 Feb 2002	YB	North-East	Bradford GHR, West Yorks	220000	17000	7.73	S	L	2002.01
430	25 Feb 2002	YB	East Mid	Ripley, Derbyshire	345000	26000	7.54	M	F	2002.01
431	25 Feb 2002	YB	North-East	Rotherham, S. Yorkshire	1060000	87500	8.25	L	F	2002.01
432	25 Feb 2002	YB	North-East	Firth Park, Sheffield	302000	23000	7.62	M	F	2002.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
433	25 Feb 2002	YB	North-West	Stockport, Cheshire	1052000	82000	7.79	L	F	2002.01
434	25 Feb 2002	Clyde	Scotland	Ellon, Aberdeenshire	354000	27050	7.64	M	Feuhold	2002.01
435	25 Feb 2002	Clyde	Scotland	Fraserburgh, Grampian	321000	25700	8.01	M	Feuhold	2002.01
436	25 Feb 2002	Clyde	Scotland	Huntly, Aberdeenshire		19400			Feuhold	2002.01
437	25 Feb 2002	Clyde	Scotland	Nairn, Moray	160000	11600	7.25	S	Feuhold	2002.01
438	25 Feb 2002	Clyde	Scotland	Peterhead, Grampian	310000	25100	8.10	M	Feuhold	2002.01
439	25 Feb 2002	Clyde	Scotland	Thurso, Highland	225000	21250	9.44	S	Feuhold	2002.01
440	25 Feb 2002	Clyde	Scotland	Turriff, Aberdeenshire	262000	20900	7.98	M	Feuhold	2002.01
441	25 Feb 2002	Clyde	Scotland	Wick, Highland	203000	18500	9.11	S	Feuhold	2002.01
442	25 Feb 2002	LTSB	South-West	Bodmin, Cornwall	252000	20000	7.92	M	F	2002.01
443	25 Feb 2002	LTSB	South-East	Braintree, Essex		42000			F	2002.01
444	25 Feb 2002	LTSB	South-East	Bromley, Kent		120000			F	2002.01
445	25 Feb 2002	LTSB	South-West	Crediton, Devon	400000	28000	7.00	M	F	2002.01
446	25 Feb 2002	LTSB	South-West	Launceston, Cornwall	275000	19000	6.91	M	F	2002.01
447	25 Feb 2002	LTSB	South-West	Okehampton, Devon	400000	27500	6.88	M	F	2002.01
448	25 Feb 2002	LTSB	South-East	Sandown, Isle of Wight	300000	22500	7.50	M	F	2002.01
449	25 Feb 2002	LTSB	North-East	Thirsk, North Yorkshire		27000			F	2002.01
450	25 Feb 2004	BW	South-West	Portishead, Bristol	365000	20000	5.48	M	F	2004.01
451	26 Nov 2002	Clyde	Scotland	Aberdeen GN, Grampian	260000	24800	9.54	M	Feuhold	2002.04
452	26 Nov 2003	LTSB	South-East	Bexhill, E. Sussex	885000	57500	6.50	L	F	2003.04
453	26 Nov 2003	LTSB	North-West	Nelson MR, Lancs	385000	27000	7.01	M	F	2003.04
454	26 Nov 2003	Other	South-East	Chichester, W. Sussex	2050000	95000	4.63	L	F	2003.04
455	27 Feb 2003	Abbey	Wales	Bridgend, Mid Glamorgan	930000	68000	7.31	L	F	2003.01
456	27 Feb 2003	BW	South-West	Bedminster, Bristol	472000	27500	5.83	M	F	2003.01
457	27 Feb 2003	BW	South-East	Bromley, Kent	878000	47500	5.41	L	F	2003.01
458	27 Feb 2003	BW	Wales	Cardiff, SMS, M. Glamorg	698000	35000	5.01	L	F	2003.01
459	27 Feb 2003	BW	South-East	Chelmsford, Essex	406000	20000	4.93	M	F	2003.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
460	27 Feb 2003	BW	South-West	Chipping Sodbury, Glos	622000	36000	5.79	L	F	2003.01
461	27 Feb 2003	BW	South-West	Downend, Bristol	604000	36000	5.96	L	F	2003.01
462	27 Feb 2003	BW	Scotland	Edinburgh CS, M.Lothian	2685000	162500	6.05	L	Feuhold	2003.01
463	27 Feb 2003	BW	South-West	Fishponds, Bristol	601000	35000	5.82	L	F	2003.01
464	27 Feb 2003	BW	South-West	Honiton, Devon	753000	47500	6.31	L	F	2003.01
465	27 Feb 2003	BW	South-East	Kingston-upon-Thames	1807000	125000	6.92	L	F	2003.01
466	27 Feb 2003	BW	South-West	Kingswood, Bristol	590000	35000	5.93	L	F	2003.01
467	27 Feb 2003	BW	London-M25	Romford, Essex	790000	48000	6.08	L	F	2003.01
468	27 Feb 2003	BW	South-West	Salisbury, Wiltshire	1720000	110000	6.40	L	F	2003.01
468	27 Feb 2003	BW	South-West	Taunton, Somerset	1755000	122856	7.00	L	F	2003.01
470	27 Feb 2003	BW	South-East	Weybridge, Surrey	603000	35000	5.80	L	F	2003.01
471	27 Feb 2003	LTSB	North-West	Ormskirk, Lancashire	470000	32500	6.91	M	F	2003.01
472	27 Jun 2001	LTSB	South-East	Bourne End, Bucks	398000	20000	5.03	M	F	2001.02
473	27 Jun 2001	LTSB	East Anglia	Cambridge, Mill Rd, Camb	401000	22500	5.61	M	F	2001.02
474	27 Jun 2001	LTSB	West Mid	Droitwich, Worcs	495000	30000	6.06	M	F	2001.02
475	27 Jun 2001	LTSB	London-M25	East Ham, London	590000	40000	6.78	L	F	2001.02
476	27 Jun 2001	LTSB	South-East	Hayes, Middlesex	410000	32500	7.93	M	F	2001.02
477	27 Jun 2001	LTSB	South-East	Henley-on-Thames, Oxon	565000	35000	6.19	L	F	2001.02
478	27 Jun 2001	LTSB	South-East	Hythe, Southampton, Hant	260000	22500	8.65	M	F	2001.02
479	27 Jun 2001	LTSB	West Mid	Kings Norton, Birmingham	280000	17500	6.25	M	F	2001.02
480	27 Jun 2001	LTSB	East Mid	Northampton, Kettering R	165000	12000	7.27	S	F	2001.02
481	27 Jun 2001	LTSB	West Mid	Leamington Spa	1870000	140000	7.49	L	F	2001.02
482	27 Jun 2001	LTSB	East Mid	Leicester SAR, Leics	342000	22500	6.58	M	F	2001.02
483	27 Jun 2001	LTSB	East Mid	Lincoln Bailgate, Lincs	375000	27000	7.20	M	F	2001.02
484	27 Jun 2001	LTSB	East Mid	Newark, Notts	762000	55000	7.22	L	F	2001.02
485	27 Jun 2001	LTSB	South-East	Pitsea, Essex	405000	25000	6.17	M	F	2001.02
486	27 Jun 2001	LTSB	London-M25	South Harrow, Middlesex	565000	42500	7.52	L	F	2001.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
487	27 Jun 2001	LTSB	London-M25	Temple Fortune, Finchley	219000	15000	6.85	S	F	2001.02
488	27 Jun 2001	LTSB	London-M25	Tooting MR, London	600000	35000	5.83	L	F	2001.02
489	27 Jun 2001	LTSB	South-West	Wotton Bassett, Wilts	326000	22500	6.90	M	F	2001.02
490	27 Jun 2001	LTSB	South-East	Ware, Hertfordshire	300000	17500	5.83	M	F	2001.02
491	27 Jun 2001	LTSB	West Mid	Aldridge, West Midlands	161000	10000	6.21	S	F	2001.02
492	27 Jun 2001	LTSB	South-East	Wealdstone, Middlesex	470000	30000	6.38	M	F	2001.02
493	27 Jun 2001	LTSB	South-West	Wellington, Somerset	231000	13500	5.84	S	F	2001.02
494	27 Nov 2001	LTSB	East Mid	Brigg, Lincolnshire	167000	12000	7.19	S	F	2001.02
495	30 Mar 2004	BW	South-West	Bideford, Devon	306000	20000	6.54	M	L	2004.01
496	30 Mar 2004	BW	South-West	Bristol, Broad Quay	365000	18000	4.93	M	F	2004.01
497	30 Mar 2004	BW	South-West	Burnham-on-Sea, Somerset	290000	15000	5.17	M	F	2004.01
498	30 Mar 2004	BW	West Mid	Evesham, Worcs	411000	20000	4.87	M	F	2004.01
499	30 Mar 2004	BW	South-West	Shirehampton, Bristol	285000	15000	5.26	M	F	2004.01
500	30 Mar 2004	BW	South-West	St. Austell, Cornwall	310000	15000	4.84	M	F	2004.01
501	30 Mar 2004	BW	Wales	Swansea, Craddock Street	430000	32000	7.44	M	L	2004.01
502	30 Mar 2004	BW	South-East	Worthing, W.Sussex	752000	33000	4.39	L	F	2004.01
503	04 Jul 2001	LTSB	South-East	Ramsgate, Kent	860000	70000	8.14	L	F	2001.03
504	17 Oct 2001	LTSB	South-East	Fleet, Hampshire	536000	40000	7.46	L	F	2001.04
505	27 Mar 2001	LTSB	South-East	Thame, Oxfordshire	437000	30000	6.86	M	F	2001.01
506	03 Dec 2002	Barclay	South-West	Bedminster, Bristol	567000	41500	7.32	L	F	2002.04
507	03 Dec 2002	LTSB	North-East	Castleford, W. Yorks		62000			F	2002.04
508	03 Dec 2002	LTSB	North-East	Goole, Yorkshire	275000	23000	8.36	M	F	2002.04
509	03 Dec 2002	LTSB	Wales	Haverford West, Pembroke	600000	44000	7.33	L	F	2002.04
510	03 Dec 2002	LTSB	Wales	Port Talbot, W.Glamorgan	317500	25000	7.87	M	F	2002.04
511	03 Dec 2002	LTSB	North-West	Runcorn, Cheshire	505000	37500	7.43	L	F	2002.04
512	03 Dec 2002	LTSB	South-East	Southampton LR, Hants	680000	48000	7.06	L	F	2002.04
513	03 Dec 2002	LTSB	North-West	New Ferry, Merseyside	229000	12500	5.46	S	F	2002.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
514	04 Feb 2002	Barclay	South-East	Maldon, Essex	762500	53350	7.00	L	F	2002.01
515	04 Feb 2002	BW	South-East	Newbury, Berkshire	485000	37500	7.73	M	F	2002.01
516	04 Feb 2002	Clyde	Scotland	Davidson's Mains, Edinbu	487500	27100	5.56	M	Feuhold	2002.01
517	04 Feb 2002	LTSB	West Mid	Bridgnorth, Shropshire	594000	42000	7.07	L	F	2002.01
518	04 Feb 2002	LTSB	South-West	Devizes, Wiltshire	550000	40000	7.27	L	F	2002.01
519	04 Feb 2002	LTSB	South-East	Farnham, Surrey	1055000	63000	5.97	L	F	2002.01
520	04 Feb 2002	LTSB	Scotland	Kirkintillock, E.Dumbart	290000	27000	9.31	M	Feuhold	2002.01
521	04 Feb 2002	LTSB	East Anglia	Lowestoft, Suffolk	750000	57500	7.67	L	F	2002.01
522	04 Feb 2002	LTSB	North-East	Scarborough, N.Yorks	905000	68500	7.57	L	F	2002.01
523	04 Feb 2002	LTSB	South-West	Tiverton, Devon	730000	54000	7.40	L	F	2002.01
524	04 Feb 2002	LTSB	South-East	Wallingford, Oxon	825000	45000	5.45	L	F	2002.01
525	06 Dec 2001	LTSB	South-West	Midsomer Norton, Somerse	365000	20000	5.48	M	F	2001.04
526	08 Oct 2003	Barclay	North-West	Crewe, Cheshire	1000000	69450	6.95	L	F	2003.04
527	08 Oct 2003	HFX	London-M25	Lewisham, London		51000			F	2003.04
528	08 Oct 2003	HFX	South-East	Tunbridge Wells, Kent		175000			F	2003.04
529	08 Oct 2003	HSBC	South-West	Boscombe, Bournemouth	460000	25000	5.43	M	F	2003.04
530	08 Oct 2003	LTSB	London-M25	Bexleyheath, Kent	1670000	115000	6.89	L	F	2003.04
531	08 Oct 2003	LTSB	South-East	Bracknell, Berkshire	1200000	80000	6.67	L	F	2003.04
532	08 Oct 2003	LTSB	North-East	Castleford, West Yorks	910000	62000	6.81	L	F	2003.04
533	08 Oct 2003	LTSB	South-East	Chatham, Kent	1080000	75000	6.94	L	F	2003.04
534	08 Oct 2003	LTSB	South-East	Epsom, Surrey	1810000	98000	5.41	L	F	2003.04
535	08 Oct 2003	LTSB	West Mid	Erdington, West Midlands		76000			F	2003.04
536	08 Oct 2003	LTSB	London-M25	Harrow, Middlesex	1220000	71500	5.86	L	F	2003.04
537	08 Oct 2003	LTSB	North-East	Kingston-upon-Hull GS	950000	65000	6.84	L	F	2003.04
538	08 Oct 2003	LTSB	North-East	Kingston-upon-Hull HR	295000	20000	6.78	M	F	2003.04
539	08 Oct 2003	LTSB	West Mid	Rugby, Warwickshire	1310000	72500	5.53	L	F	2003.04
540	08 Oct 2003	LTSB	North-East	Seaham, County Durham	132500	7500	5.66	S	F	2003.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
541	08 Oct 2003	LTSB	West Mid	Solihull, Poplar Rd, WM		93500			F	2003.04
542	09 Oct 2002	Abbey	South-East	Luton, Beds	1250000	85000	6.80	L	F	2002.04
543	09 Oct 2002	COOP	North-East	Rotherham, S. Yorks	865000	57000	6.59	L	F	2002.04
544	09 Oct 2002	HSBC	South-East	Caterham CR, Surrey	255000	14000	5.49	M	F	2002.04
545	09 Oct 2002	HSBC	East Anglia	Thetford, Norfolk	450000	23000	7.33	M	F	2002.04
546	09 Oct 2002	LTSB	South-West	Bishopsworth, Bristol	290000	19000	6.55	M	F	2002.04
547	09 Oct 2002	LTSB	South-West	Bodmin, Cornwall	275000	20000	7.27	M	F	2002.04
548	09 Oct 2002	LTSB	South-East	Broadstairs, Kent	338000	21500	6.36	M	F	2002.04
449	09 Oct 2002	LTSB	North-East	Rotherham, S. Yorks	545000	42500	7.80	L	F	2002.04
550	09 Oct 2002	LTSB	West Mid	Wednesfield, Wolverhampt	480000	30000	6.25	M	F	2002.04
551	09 Oct 2002	Nationw	North-West	Barrow-in-Furness, Cumbr	407500	31000	7.61	M	F	2002.04
552	09 Oct 2002	YB	North-East	Firth Park, Sheffield	325000	23000	7.08	M	F	2002.04
553	10 Feb 2004	Abbey	South-West	Weston-super-Mare, Somer	240000	14250	5.94	S	F	2004.01
554	10 Feb 2004	Barclay	South-East	Dorking, Surrey	1710000	94000	5.50	L	F	2004.01
555	10 Feb 2004	Barclay	South-East	South Lancing, W.Sussex	620000	37900	6.11	L	F	2004.01
556	10 Feb 2004	COOP	South-West	Gloucester, Glos	800000	50000	6.25	L	F	2004.01
557	10 Feb 2004	HSBC	South-East	Chatham, Kent	1000000	70000	7.00	L	F	2004.01
558	10 Feb 2004	LTSB	West Mid	Evesham, Worcs	1115000	60000	5.38	L	F	2004.01
559	10 Feb 2004	LTSB	North-West	Liverpool GCS, Merseysid	1500000	110000	7.33	L	L	2004.01
560	10 Feb 2004	LTSB	South-East	Winton, Bournemouth, Dor	600000	35000	5.83	L	F	2004.01
561	11 Feb 2003	Abbey	South-East	Godalming, Surrey	520000	32500	6.25	L	F	2003.01
562	11 Feb 2003	LTSB	North-West	Altrincham, Cheshire	865000	60000	6.94	L	F	2003.01
563	11 Feb 2003	LTSB	East Mid	Loughborough, Leics	1240000	96500	7.78	L	F	2003.01
564	11 Feb 2003	LTSB	North-East	South Shields, Tyne & We	955000	73500	7.70	L	F	2003.01
565	11 Feb 2003	NatWest	East Anglia	North Walsham, Norfolk		11000			F	2003.01
566	11 Jul 2001	Abbey	East Anglia	Lowestoft, Suffolk	512000	37000	7.23	L	F	2001.03
567	11 Jul 2001	Barclay	North-East	Sunderland, Tyne & Wear	852500	64000	7.51	L	F	2001.03

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
568	11 Jul 2001	HSBC	South-West	Coleford, Glos		9000			F	2001.03
569	11 Jul 2001	HSBC	South-East	Woolston, Southampton, H	287000	22500	7.84	M	F	2001.03
570	11 Jul 2001	LTSB	South-East	Sidcup, Kent	465000	35000	7.53	M	F	2001.03
571	11 Jul 2001	NatWest	London-M25	Plaistow, London		16000			F	2001.03
572	11 Jul 2002	COOP	North-West	Didsbury, Gt. Manchester	575000	32000	5.57	L	F	2002.03
573	11 Jul 2002	HSBC	Wales	Merthyr Tydfil, M.Glamor	995000	79000	7.94	L	F	2002.03
574	11 Jul 2002	HSBC	South-East	Watford, Herttfordshire		120000			L	2002.03
575	11 Jul 2002	LTSB	North-West	Blackpool, Lancashire	900000	60000	6.67	L	F	2002.03
576	11 Jul 2002	LTSB	South-East	Bromley, Kent	1650000	120000	7.27	L	F	2002.03
577	11 Jul 2002	LTSB	North-East	Castleford, W. Yorks		62000			F	2002.03
578	11 Jul 2002	LTSB	East Anglia	Gorleston-on-Sea, Norfol	375000	27500	7.33	M	F	2002.03
579	11 Jul 2002	LTSB	North-West	Prescot, Merseyside	255000	19000	7.45	M	F	2002.03
580	11 Jul 2002	LTSB	West Mid	Tunstall, Stoke-on-Trent	182000	12500	6.87	S	F	2002.03
581	11 Jul 2002	Nationw	North-East	Wakefield, West Yorks	502500	39000	7.76	L	F	2002.03
582	11 Jul 2002	RBS	South-East	Maidenhead, Berkshire		200000			F	2002.03
583	11 Jul 2002	YB	North-West	Stalybridge, Gt.Manchest	162000	10000	6.17	S	F	2002.03
584	11 Jul 2002	YB	North-East	Wombwell, South Yorks		21000			F	2002.03
585	17 Oct 2001	Abbey	South-East	Andover, Hampshire	760000	53100	6.99	L	F	2001.04
586	17 Oct 2001	Abbey	South-East	Haywards Heath, Sussex	660000	52500	7.95	L	F	2001.04
587	17 Oct 2001	LTSB	North-West	Atherton, Gt. Manchester	235000	52500	8.51	S	F	2001.04
588	17 Oct 2001	LTSB	South-East	Bedford, Beds	1045000	87500	8.37	L	F	2001.04
589	17 Oct 2001	LTSB	East Mid	Brackley, Northants	322500	22500	6.98	M	F	2001.04
590	17 Oct 2001	LTSB	Wales	Cardiff WR		35000			F	2001.04
591	17 Oct 2001	BW	South-West	Clifton, Bristol	1355000	97000	7.16	L	F	2001.04
592	17 Oct 2001	LTSB	West Mid	Coventry, Warks	1675000	137000	8.18	L	F	2001.04
593	17 Oct 2001	LTSB	North-East	Dewsbury, West Yorkshire	315000	28000	8.89	M	F	2001.04
594	17 Oct 2001	LTSB	South-East	Fleet, Hampshire	536000	40000	7.46	L	F	2001.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
595	17 Oct 2001	LTSB	East Anglia	Halesworth, Suffolk	210000	12500	5.95	S	F	2001.04
596	17 Oct 2001	LTSB	Wales	Holyhead, Anglesey	160000	12500	7.81	S	F	2001.04
597	17 Oct 2001	LTSB	Wales	Lampeter, Ceredigion	240000	16000	6.67	S	F	2001.04
598	17 Oct 2001	LTSB	South-East	Lewes, East Sussex	455000	31500	6.92	M	F	2001.04
599	17 Oct 2001	LTSB	West Mid	Nuneaton, Warks	590000	40000	6.78	L	F	2001.04
600	17 Oct 2001	LTSB	Wales	Penarth, S. Glamorgan	367500	27000	7.35	M	F	2001.04
601	17 Oct 2001	LTSB	Wales	Pontypridd, M. Glamorgan		30000			F	2001.04
602	17 Oct 2001	LTSB	South-East	Royston, Hertfordshire	453000	35500	7.84	M	F	2001.04
603	17 Oct 2001	LTSB	East Mid	Sherwood, Notts		25000			F	2001.04
604	17 Oct 2001	LTSB	South-East	Seaford, East Sussex	315000	22600	7.17	M	F	2001.04
605	17 Oct 2001	LTSB	South-East	Thorpe Bay, Essex	312500	16500	5.28	M	F	2001.04
606	17 Oct 2001	LTSB	South-West	Wareham, Dorset	480000	35000	7.29	M	F	2001.04
607	17 Oct 2001	NatWest	Wales	Penarth, S. Glamorgan	188000	11000	5.85	S	F	2001.04
608	19 Mar 2002	LTSB	South-East	Stanford-le-Hope, Essex	282500	20000	7.08	M	F	2002.01
609	19 Mar 2002	Abbey	Wales	Caernarfon, Gwynedd	147000	11250	7.65	S	F	2002.01
610	19 Mar 2002	Abbey	South-East	Gravesend, Kent	515000	30000	5.83	L	F	2002.01
611	19 Mar 2002	Abbey	North-East	Harrogate, North Yorks	1170000	75000	6.41	L	F	2002.01
612	19 Mar 2002	Abbey	South-East	Worcester Park, Surrey	272000	15750	5.79	M	F	2002.01
613	19 Mar 2002	BW	South-West	Barnstable, Devon	2310000	171040	7.40	L	F	2002.01
614	19 Mar 2002	BW	South-West	Penzance, Cornwall	1550000	116388	7.51	L	F	2002.01
615	19 Mar 2002	HFX	London-M25	Croydon, Surrey	1810000	123000	6.80	L	F	2002.01
616	19 Mar 2002	LTSB	East Anglia	Aylsham, Norfolk	260000	20000	7.69	M	F	2002.01
617	19 Mar 2002	LTSB	North-East	Byker, South Shields	251000	19800	7.89	M	F	2002.01
618	19 Mar 2002	LTSB	South-East	East Wittering, W.Sussex	196000	14000	7.14	S	F	2002.01
619	19 Mar 2002	LTSB	South-West	Helston, Cornwall	280000	22000	7.86	M	F	2002.01
620	19 Mar 2002	LTSB	South-East	Lewes, East Sussex	500000	34500	6.90	M	F	2002.01
621	19 Mar 2002	LTSB	Wales	Llandeillo, Carmarthensh	180000	14500	8.06	S	F	2002.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
622	19 Mar 2002	LTSB	Wales	Llantwit Major, V of Gla	160000	11500	7.19	S	F	2002.01
623	19 Mar 2002	LTSB	Wales	Newcastle Emlyn, Ceredig		19000			F	2002.01
624	19 Mar 2002	LTSB	South-East	Rottingdean, Brighton &	411000	22650	5.51	M	F	2002.01
625	19 Mar 2002	LTSB	South-West	St. Austell, Cornwall	377500	29000	7.68	M	F	2002.01
626	19 Mar 2002	LTSB	South-East	Tadley, Hampshire	195000	15000	7.69	S	F	2002.01
627	19 Mar 2002	LTSB	West Mid	Tettenhall. West Mid	431000	29000	6.73	M	F	2002.01
628	19 Mar 2002	LTSB	West Mid	Shrewsbury, Shropshire		174370			F	2002.01
629	19 May 2004	Barclay	South-East	Purley, Surrey	580000	30000	5.17	L	F	2004.02
630	19 May 2004	Barclay	South-East	Stopsley, Luton, Beds	500000	28800	5.76	M	F	2004.02
631	19 May 2004	Barclay	North-West	Wallasey, Liscard, Merse		72500			F	2004.02
632	19 May 2004	BW	South-West	Penzance, Cornwall		116388			F	2004.02
633	19 May 2004	BW	South-East	Southend-on-Sea, Essex	530000	32500	6.13	L	L	2004.02
634	19 May 2004	Clyde	Scotland	Glasgow, Duke St	300000	17500	5.83	M	Feuhold	2004.02
635	19 May 2004	LTSB	South-East	Coulsdon, Surrey	545000	28000	5.14	L	F	2004.02
636	19 May 2004	LTSB	Scotland	Hamilton, S. Lanarkshire	665000	50000	7.52	L	Feudal	2004.02
637	19 May 2004	LTSB	East Mid	Leicester AR, Leics	465000	22500	4.84	M	F	2004.02
638	19 May 2004	LTSB	East Mid	Melton Mowbray, Leics	732500	42500	5.80	L	F	2004.02
639	19 May 2004	Other	West Mid	Erdington, W. Midlands	600000	35500	5.92	L	F	2004.02
640	19 May 2004	Other	South-West	Plymouth AW, Devon		64000			L	2004.02
641	21 Jun 2001	LTSB	Scotland	Corstorphine, Edinburgh	491000	38000	7.74	M	Feuhold	2001.02
642	21 Jun 2001	LTSB	Scotland	Renfrew, Renfrewshire	250000	23500	9.40	S	Feuhold	2001.02
643	21 May 2003	BW	South-West	Bath, Somerset	1470000	74680	5.08	L	F	2003.02
644	21 May 2003	BW	South-East	Maidenhead, Berks	685000	43128	6.30	L	F	2003.02
645	21 May 2003	LTSB	Wales	Ystrad Mynach, Hengoed	212000	15000	7.08	S	F	2003.02
646	21 May 2003	LTSB	South-East	Horsham, W. Sussex	3600000	225000	6.25	L	F	2003.02
647	21 May 2003	LTSB	North-East	Meltham, West Yorks	69000	4400	6.38	S	F	2003.02
648	21 May 2003	LTSB	North-West	Nelson MR, Lancashire	370000	27000	7.30	M	F	2003.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
649	21 May 2003	NatWest	South-East	Hitchin, Hertfordshire	470000	20825	4.43	M	F	2003.02
650	21 May 2003	NatWest	East Anglia	North Walsham, Norfolk	151000	11000	7.28	S	F	2003.02
651	21 May 2003	YB	West Mid	Bromsgrove, W. Midlands	765000	50000	6.54	L	F	2003.02
652	22 May 2001	HSBC	West Mid	Bromyard, Herefordshire	135000	10000	7.41	S	F	2001.02
653	22 May 2001	HSBC	North-East	Cleckheaton, West Yorks	195000	15500	7.95	S	F	2001.02
654	22 May 2001	HSBC	South-East	Coulsdon, Surrey	262500	16000	6.10	M	F	2001.02
655	22 May 2001	HSBC	South-West	Crediton, Devon	200000	15000	7.50	S	F	2001.02
656	22 May 2001	HSBC	South-West	Dursley, Glos	200000	15000	7.50	S	F	2001.02
657	22 May 2001	HSBC	North-East	Hornsea, East Riding Yks	127000	9500	7.48	S	F	2001.02
658	22 May 2001	HSBC	Wales	Llanfair Caereilion, Pow	140000	11500	8.21	S	F	2001.02
659	22 May 2001	HSBC	Wales	Llanfyllin, Powys	74000	5000	7.43	S	F	2001.02
660	22 May 2001	HSBC	South-East	Shirley, Southampton	280000	20500	7.32	M	F	2001.02
661	22 May 2001	HSBC	South-East	Tadley, Hampshire	252500	17000	6.73	M	F	2001.02
662	22 May 2001	LTSB	Wales	Tonyrefail, M.Glamorgan	105000	9000	8.57	S	F	2001.02
663	22 May 2001	YB	North-East	Batley, West Yorkshire	275000	17000	6.18	M	F	2001.02
664	23 May 2002	Abbey	East Mid	Grimsby, Lincolnshire	1150000	75000	6.52	L	F	2002.02
665	23 May 2002	BW	South-West	Bridgwater, Somerset	630000	41888	6.65	L	F	2002.02
666	23 May 2002	BW	South-West	Bridport, Dorset	920000	68660	7.46	L	F	2002.02
667	23 May 2002	Clyde	Scotland	Laurencekirk, Aberdeensh	140000	9600	6.86	S	Heritabl	2002.02
668	23 May 2002	HFC	North-East	Kingston-upon-Hull GS	219000	17500	7.99	S	F	2002.02
669	23 May 2002	LTSB	South-East	Alton, Hampshire	715000	41000	5.73	L	F	2002.02
670	23 May 2002	LTSB	West Mid	Blackheath, W. Midlands	281000	18500	6.58	M	F	2002.02
671	23 May 2002	LTSB	Wales	Caernarfon, Gwynedd	330000	27000	8.18	M	F	2002.02
672	23 May 2002	LTSB	London-M25	Cheam, Surrey	428000	24500	5.72	M	F	2002.02
673	23 May 2002	LTSB	South-East	Herne Bay, Kent	502500	35000	6.97	L	F	2002.02
674	23 May 2002	LTSB	East Mid	Shirebrook, Derbyshire	192000	12500	6.51	S	F	2002.02
675	23 May 2002	LTSB	East Anglia	St. Neotts, Cambs	831000	56500	6.80	L	F	2002.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
676	23 May 2002	LTSB	North-West	Stockport Hazel Grove		40000			F	2002.02
677	23 May 2002	LTSB	South-West	Tewkesbury, Glos	500000	31000	6.20	M	F	2002.02
678	23 May 2002	LTSB	South-West	Warminster, Wilts	605000	37500	6.20	L	F	2002.02
679	23 May 2002	LTSB	South-West	Winchcombe, Glos	382000	22500	5.89	M	F	2002.02
680	23 May 2002	NatWest	North-East	Firth Park, sheffield	215000	17300	8.05	S	F	2002.02
681	23 May 2002	Coventr	South-East	Banbury, Oxfordshire	975000	70000	7.18	L	F	2002.02
682	23 May 2002	Other	West Mid	Shrewsbury, Shropshire	640000	45000	7.03	L	F	2002.02
683	23 May 2002	YB	South-East	Luton, Bedfordshire		82250			L	2002.02
684	25 Mar 2003	Abbey	Wales	Caernarfon, Gwynedd	161000	11250	6.99	S	F	2003.01
685	25 Mar 2003	LTSB	East Mid	Grantham, Lincs		50000				2003.01
686	25 Mar 2003	LTSB	South-West	Wareham, Dorset	540000	35000	6.48	L	F	2003.01
687	25 Mar 2004	Nationw	South-East	Burgess Hill, W. Sussex	410000	16450	4.01	M		2004.01
688	25 Mar 2004	YB	North-West	Stockport, Cheshire	1530000	82000	5.36	L	F	2004.01
689	27 Mar 2001	LTSB	South-East	Wantage, Oxfordshire	495000	41000	8.28	M	F	2001.01
690	03 Apr 2001	Abbey	London-M25	Peckham, London	589000	50000	8.49	L	F	2001.02
691	03 Apr 2001	LTSB	Wales	Merthyr Tidfyl, M.Glam	856000	72000	8.41	L	F	2001.02
692	27 Feb 2001	LTSB	South-East	Beckenham, Kent	310000	28100	9.06	M	F	2001.01
693	27 Jun 2001	LTSB	Scotland	Aberdeen DP, Grampian	1210000	106000	8.76	L	Feuhold	2001.02
694	27 Jun 2001	LTSB	South-East	Amesbury, Wiltshire	370000	32500	8.78	M	F	2001.02
695	27 Jun 2001	LTSB	Wales	Ammanford, Dyfed	276000	22500	8.15	M	F	2001.02
696	27 Jun 2001	LTSB	Scotland	Aberdeen AP, Grampian	860000	82600	9.60	L	Feuhold	2001.02
697	27 Jun 2001	LTSB	South-East	Ashford, Kent	1295000	100000	7.72	L	F	2001.02
698	27 Jun 2001	LTSB	South-East	Ashford A, Middlesex	288000	22500	7.81	M	F	2001.02
699	27 Jun 2001	LTSB	East Mid	Atherstone, Warks	252000	17500	6.94	M	F	2001.02
700	27 Jun 2001	LTSB	North-East	Bishop Auckland, C.Durha	330000	25000	7.58	M	F	2001.02
701	27 Jun 2001	LTSB	East Mid	Beeston, Notts	396000	35000	8.84	M	F	2001.02
702	27 Jun 2001	LTSB	South-East	Bicester, Oxfordshire	500000	44000	8.80	M	F	2001.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
703	27 Jun 2001	LTSB	West Mid	Bloxwich, W. Midlands	301000	22500	7.48	M	F	2001.02
704	27 Jun 2001	LTSB	West Mid	Bradmore, Wolverhampton	135000	10000	7.41	S	F	2001.02
705	27 Jun 2001	LTSB	South-West	Brixham, Devon	310000	25000	8.06	M	F	2001.02
706	27 Jun 2001	LTSB	Scotland	Bucksburn, Aberdeen	187500	17500	9.33	S	Feuhold	2001.02
707	27 Jun 2001	LTSB	South-West	Bude, Cornwall	221000	12500	5.66	S	F	2001.02
708	27 Jun 2001	LTSB	North-West	Burslem, Stoke-on-Trent	160000	10000	6.25	S	F	2001.02
709	27 Jun 2001	LTSB	South-West	Calne, Wiltshire	355000	25000	4.23	M	F	2001.02
710	27 Jun 2001	LTSB	South-West	Camborne, Cornwall	175000	11000	6.29	S	F	2001.02
711	27 Jun 2001	LTSB	Wales	Cardigan, Dyfed	265000	22500	8.49	M	F	2001.02
712	27 Jun 2001	LTSB	North-West	Cheadle, Gt. Manchester	330000	30000	9.09	M	F	2001.02
713	27 Jun 2001	LTSB	East Mid	Coalville, Leics	460000	37500	8.15	M	F	2001.02
714	27 Jun 2001	LTSB	South-East	Corringham, Essex	191000	17500	9.16	S	F	2001.02
715	27 Jun 2001	LTSB	West Mid	Cotteridge, Birmingham	241000	17500	7.26	S	F	2001.02
716	27 Jun 2001	LTSB	West Mid	Cradley Heath, Birmingha	210000	18750	8.93	S	F	2001.02
717	27 Jun 2001	LTSB	South-West	Crownhill, Plymouth	115000	9500	8.26	S	F	2001.02
718	27 Jun 2001	LTSB	Scotland	Aberdeen CS, Grampian	255000	20000	7.84	M	Feuhold	2001.02
719	27 Jun 2001	LTSB	East Mid	Derby OR, Derbyshire	260000	20000	7.69	M	F	2001.02
720	27 Jun 2001	LTSB	Scotland	Dyce, Aberdeen, Grampian	300000	25000	8.33	M	F	2001.02
721	27 Jun 2001	LTSB	East Anglia	Felixstowe, Suffolk	325000	23000	7.08	M	F	2001.02
722	27 Jun 2001	LTSB	East Anglia	Great Yarmouth, Norfolk	560000	47500	8.48	L	L	2001.02
723	27 Jun 2001	LTSB	London-M25	Gants Hill. Ilford	329000	20000	6.08	M	F	2001.02
724	27 Jun 2001	LTSB	West Mid	Great Bridge, Birmingham	171000	15000	8.77	S	F	2001.02
725	27 Jun 2001	LTSB	West Mid	Harborne, Birmingham	842000	57500	6.83	L	F	2001.02
726	27 Jun 2001	LTSB	South-East	Hatfield, Hertfordshire	273000	20000	7.33	M	F	2001.02
727	27 Jun 2001	LTSB	South-East	Hazlemere, High Wycombe	191000	10000	5.24	S	F	2001.02
728	27 Jun 2001	LTSB	South-East	Hoddesdon, Hertfordshire	386000	23000	5.96	M	F	2001.02
739	27 Jun 2001	LTSB	North-East	Kingston-upon-Hull GS	710000	65000	9.15	L	F	2001.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
730	27 Jun 2001	LTSB	North-East	Leeds AS, West Yorks	2855000	192500	6.74	L	F	2001.02
731	27 Jun 2001	LTSB	East Mid	Leicester MR, Leics	200000	12500	6.25	S	F	2001.02
732	27 Jun 2001	LTSB	East Mid	Lincoln, Lincs	745000	55000	7.38	L	F	2001.02
733	27 Jun 2001	LTSB	West Mid	Longbridge, Birmingham	150000	10000	6.67	S	F	2001.02
734	27 Jun 2001	LTSB	East Mid	Louth, Lincolnshire	364000	28500	7.83	M	F	2001.02
735	27 Jun 2001	LTSB	South-East	Malmesbury, Wilts	255000	17500	6.86	M	F	2001.02
736	27 Jun 2001	LTSB	East Mid	Melton Mowbray, Leics	491000	42500	8.66	M	F	2001.02
737	27 Jun 2001	LTSB	London-M25	Mill Hill, NW7	370000	30000	8.11	M	F	2001.02
738	27 Jun 2001	LTSB	Wales	Morriston, Swansea	232000	18500	7.97	S	F	2001.02
739	27 Jun 2001	LTSB	Wales	Maindee, Newport, Gwent	250000	20000	8.00	S	F	2001.02
740	27 Jun 2001	LTSB	East Mid	Nottingham, Notts	1460000	100000	6.85	L	F	2001.02
741	27 Jun 2001	LTSB	West Mid	Oldbury, West Midlands	300000	27000	9.00	M	F	2001.02
742	27 Jun 2001	LTSB	North-West	Ormskirk, Lancashire	352000	32500	9.23	M	F	2001.02
743	27 Jun 2001	LTSB	West Mid	Oxley, Wolverhampton	193000	16500	8.55	S	F	2001.02
744	27 Jun 2001	LTSB	East Anglia	Peterborough, Cambs	640000	50000	7.81	L	F	2001.02
745	27 Jun 2001	LTSB	Wales	Pontypool, Gwent	152000	10000	6.58	S	F	2001.02
746	27 Jun 2001	LTSB	Wales	Port Talbot, Neath		25000			F	2001.02
747	27 Jun 2001	LTSB	West Mid	Quinton, Birmingham	236000	20000	8.47	S	F	2001.02
748	27 Jun 2001	LTSB	Wales	Rumney, Cardiff	127000	10000	7.87	S	F	2001.02
749	27 Jun 2001	LTSB	West Mid	Selly Oak, Birmingham	232000	17500	7.54	S	F	2001.02
750	27 Jun 2001	LTSB	South-West	Sherbourne, Dorset	473000	30000	6.34	M	F	2001.02
751	27 Jun 2001	LTSB	South-East	Shirley, Southampton	375000	30000	8.00	M	F	2001.02
752	27 Jun 2001	LTSB	South-East	Worcester SJ, Worcs	311000	20000	6.43	M	F	2001.02
752	27 Jun 2001	LTSB	Scotland	Aberdeen, St.Machar, Gra	310000	25000	8.06	M	F	2001.02
754	27 Jun 2001	LTSB	South-West	Torquay, St.Marychurch	262000	15000	5.73	M	F	2001.02
755	27 Jun 2001	LTSB	North-East	Stockton-on-Tees, Clevel	531000	45000	8.47	L	L	2001.02
756	27 Jun 2001	LTSB	West Mid	Stone, Staffs	275000	21000	7.64	M	F	2001.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
757	27 Jun 2001	LTSB	North-East	Sunderland, Tyne & Wear	440000	40000	9.09	M	F	2001.02
758	27 Jun 2001	LTSB	North-West	West Kirby, Merseyside	251000	17500	6.97	M	F	2001.02
759	27 Jun 2001	LTSB	West Mid	Ward End, Birmingham	407000	32500	7.99	M	F	2001.02
760	27 Jun 2001	LTSB	West Mid	Wednesbury, Birmingham	156000	10000	6.41	S	F	2001.02
761	27 Jun 2001	LTSB	South-East	Welling, Kent	247000	17500	7.09	S	F	2001.02
762	27 Jun 2001	LTSB	London-M25	Wembley, Middlesex	511000	37500	7.34	L	F	2001.02
763	27 Jun 2001	LTSB	South-West	Westbourne, Dorset	630000	47500	7.54	L	F	2001.02
764	27 Jun 2001	LTSB	Wales	Whitchurch, Cardiff	136000	10000	7.35	S	F	2001.02
765	27 Jun 2001	LTSB	South-West	Wimbourne, Dorset	573000	40000	6.98	L	F	2001.02
766	31 May 2001	LTSB	Wales	Barry, Mid Glamorgan	490000	48500	9.90	M	F	2001.02
767	08 Dec 2005	HFC	East Mid	Scunthorpe, Lincs	490000	28000	5.71	M	F	2005.04
768	08 Dec 2005	LTSB	London-M25	Hounslow, Middlesex		72000			F	2005.04
769	08 Dec 2005	RBS	South-East	St. Albans, Herts	2500000	132000	5.28	L	F	2005.04
770	08 Dec 2005	HFX	London-M25	Hounslow, Middlesex		90000			F	2005.04
771	08 Dec 2005	C+G	South-East	Witney, Oxfordshire	1650000	82500	5.00	L	F	2005.04
772	08 Dec 2005	C+G	North-East	Middlesbrough, Cleveland	655000	28000	4.27	L	F	2005.04
773	07 Feb 2006	LTSB	East Mid	Shirebrook, Derbyshire	260000	12500	4.62	M	F	2006.01
774	07 Feb 2006	HFX	South-East	Dover, Kent	1250000	68000	5.44	L	F	2006.01
775	07 Feb 2006	Woolwic	North-West	Wigan, Lancashire	1610000	68650	4.26	L	F	2006.01
776	07 Feb 2006	LTSB	South-East	Maldon, Essex	573000	22000	3.84	L	F	2006.01
777	07 Feb 2006	HSBC	South-East	Great Dunmow, Essex	680000	37500	5.51	L	F	2006.01
778	22 Mar 2006	Barclay	North-West	Chorlton cum Hardy, GMan	837000	46580	5.60	L	F	2006.01
779	22 Mar 2006	Barclay	North-East	Chester-le-Street, C. Du	930000	40000	4.30	L	F	2006.01
780	22 Mar 2006	Barclay	North-West	Penrith, Cumbria	852500	36000	4.22	L	F	2006.01
781	22 Mar 2006	Barclay	South-East	Hatfield, Herts	760000	36000	4.74	L	F	2006.01
782	22 Mar 2006	C+G	South-West	Taunton, Somerset	1340000	64000	4.78	L	F	2006.01
783	22 Mar 2006	Barclay	South-East	Stansted, Essex	690000	25000	3.62	L	F	2006.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
784	22 Mar 2006	Barclay	South-East	Clacton-on-Sea, Essex	640000	27500	4.30	L	F	2006.01
785	22 Mar 2006	Barclay	East Anglia	Fakenham, Norfolk	790000	32500	4.11	L	F	2006.01
786	22 Mar 2006	Barclay	South-East	Woodstock, Oxfordshire	480000	19000	3.96	M	F	2006.01
787	22 Mar 2006	Barclay	East Anglia	Watton, Norfolk	575000	25000	4.35	L	F	2006.01
788	22 Mar 2006	Other	South-East	Bedford, Beds	1040000	52500	5.05	L	F	2006.01
789	22 Mar 2006	Barclay	North-West	Wellington, Shropshire	700000	40000	5.71	L	F	2006.01
790	22 Mar 2006	Barclay	East Anglia	Ramsey, Cambridgeshire	457000	25000	5.46	M	L	2006.01
791	22 Mar 2006	Nationw	London-M25	Golders Green, NW11	865000	40000	4.62	L	F	2006.01
792	22 Mar 2006	LTSB	South-East	Stanford-le-Hope, Essex	500000	29000	5.80	M	F	2006.01
793	22 Mar 2006	Barclay	London-M25	Whitton, Middlesex	500000	28850	5.17	M	F	2006.01
794	22 Mar 2006	Barclay	West Mid	Knowle, Solihull	510000	25000	4.90	L	F	2006.01
795	16 May 2006	Barclay	South-East	Chertsey, Surrey	550000	29000	5.27	L	F	2006.02
796	16 May 2006	Barclay	London-M25	Thornton Heath, Croydon	800000	42500	5.31	L	F	2006.02
797	16 May 2006	Barclay	East Anglia	Aylsham, Norfolk	770000	30000	3.90	L	F	2006.02
798	16 May 2006	Barclay	South-East	Thorpe Bay, Essex	365000	14000	3.84	M	F	2006.02
799	16 May 2006	Barclay	East Mid	Leicester, Narborough Rd	805000	27000	3.35	L	F	2006.02
800	16 May 2006	Barclay	South-East	Borehamwood, Herts	1210000	47500	3.93	L	F	2006.02
801	16 May 2006	Barclay	East Anglia	Old Fletton, Peterboroug	870000	36500	4.20	L	F	2006.02
802	16 May 2006	Barclay	East Anglia	Gorleston, Norfolk	680000	32500	4.78	L	F	2006.02
803	16 May 2006	Barclay	East Mid	Luton DR, Beds	805000	27500	3.42	L	F	2006.02
804	16 May 2006	Barclay	South-East	Maldon, Essex	1220000	61500	5.04	L	F	2006.02
805	16 May 2006	LTSB	North-West	Runcorn, Cheshire	680000	37500	5.51	L	F	2006.02
806	16 May 2006	Barclay	East Anglia	Wymondham, Norfolk	785000	35750	4.55	L	F	2006.02
807	16 May 2006	Barclay	North-West	Bury, Lancashire	1190000	61500	5.17	L	F	2006.02
808	16 May 2006	Barclay	North-East	Sheffield, Glossop Rd	1000000	44100	4.41	L	F	2006.02
809	16 May 2006	Barclay	South-West	Brixham, Devon	535000	22000	4.11	L	F	2006.02
810	16 May 2006	Barclay	North-West	Morecambe, Lancs	670000	29000	4.33	L	F	2006.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
811	16 May 2006	Barclay	East Mid	Allenton, Derbyshire	540000	25000	4.63	L	F	2006.02
812	16 May 2006	Barclay	East Mid	Allertree, Derbyshire	555000	27000	4.86	L	F	2006.02
813	16 May 2006	Barclay	South-West	Plymouth, Matley Plain	1040000	45700	4.39	L	F	2006.02
814	16 May 2006	Barclay	South-West	Paignton, Devon	565000	30100	5.33	L	F	2006.02
815	16 May 2006	Barclay	North-West	Keswick, Cumbria	865000	49000	5.66	L	L	2006.02
816	16 May 2006	LTSB	South-West	Penzance, Cornwall		82000			F	2006.02
817	16 May 2006	Barclay	North-West	Liverpool, Prescot Rd	910000	46250	5.08	L	F	2006.02
818	16 May 2006	Barclay	South-East	Aldershot, Hampshire	750000	39500	5.27	L	F	2006.02
819	16 May 2006	LTSB	West Mid	Wombourne, Wolverhampton	730000	37500	5.14	L	F	2006.02
820	05 Jul 2006	HSBC	West Mid	Knowle, West Midlands	600000	30000	5.00	L	F	2006.03
821	05 Jul 2006	HSBC	South-East	Eastleigh, Hampshire		35000			F	2006.03
822	05 Jul 2006	HSBC	South-East	Brighton, E.Ditching Rd	605000	22500	3.72	L	F	2006.03
823	05 Jul 2006	Barclay	West Mid	Sutton Coldfield, W. Mid	540000	47025	4.63	L	F	2006.03
824	05 Jul 2006	Woolwic	West Mid	Sutton Coldfield, W. Mid	465000	22000	4.73	M	F	2006.03
825	05 Jul 2006	Barclay	South-East	Goring-on-Sea, W.Sussex	345000	15600	4.52	M	F	2006.03
826	05 Jul 2006	HSBC	South-East	Dunstable, Beds	630000	32500	5.16	L	F	2006.03
827	05 Jul 2006	HSBC	South-East	Gillingham, Kent	655000	33000	5.04	L	F	2006.03
828	05 Jul 2006	Barclay	South-East	Beaconsfield, Bucks	900000	30000	3.33	L	F	2006.03
829	05 Jul 2006	HSBC	North-East	Driffield, E. Yorks	480000	21500	4.48	M	F	2006.03
830	05 Jul 2006	HSBC	South-East	Camberley, Surrey	1250000	50000	4.00	L	F	2006.03
831	05 Jul 2006	HSBC	South-East	Farnham, surrey	1700000	90000	5.29	L	F	2006.03
832	05 Jul 2006	HSBC	East Mid	Skegness, Lincolnshire	955000	47500	4.97	L	F	2006.03
833	05 Jul 2006	HSBC	East Mid	Beeston, notts	1600000	90000	5.63	L	F	2006.03
834	05 Jul 2006	HSBC	West Mid	Harborne, Birmingham	1490000	67500	4.53	L	F	2006.03
835	05 Jul 2006	HSBC	North-East	Leeds, Blenheim T	565000	25000	4.42	L	F	2006.03
836	05 Jul 2006	HSBC	East Mid	Ripley, Derbyshire	505000	25000	4.95	L	F	2006.03
837	05 Jul 2006	HSBC	Wales	Pontypridd	1200000	65000	5.42	L	F	2006.03

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
838	05 Jul 2006	Barclay	South-East	Strood, Kent	727500	35300	4.85	L	F	2006.03
839	05 Jul 2006	Barclay	South-East	Didcot, Oxfordshire	770000	45500	5.91	L	F	2006.03
840	05 Jul 2006	HSBC	North-East	Middlesbrough AR, Clevel	1650000	92000	5.58	L	F	2006.03
841	05 Jul 2006	HSBC	North-West	Congleton, Cheshire	510000	26000	5.10	L	F	2006.03
842	05 Jul 2006	HSBC	West Mid	Sutton Coldfield, WM	840000	35000	4.17	L	F	2006.03
843	05 Jul 2006	HSBC	East Mid	Bulwell, Notts	690000	35000	5.07	L	F	2006.03
844	05 Jul 2006	Barclay	North-West	Didsbury, Gt Manchester	1075000	56000	5.21	L	F	2006.03
845	05 Jul 2006	HSBC	North-West	Formby, Lancashire	700000	33500	4.79	L	F	2006.03
846	05 Jul 2006	HSBC	North-West	Hyde, Gt.Manchester	610000	30000	4.92	L	F	2006.03
847	05 Jul 2006	LTSB	North-West	Reddish, Cheshire	240000	12900	5.38	S	L	2006.03
848	22 Feb 2001	Other	Scotland	Dumfries, Dumfries & Gal	530000	38800	7.32	L	Feudal	2001.01
849	22 Feb 2001	LTSB	Scotland	Dumfries, Dumfries & Gal	887000	75000	8.46	L	Feudal	2001.01
850	22 May 2002	BW	South-West	Gloucester HL, Glos	1195000	88372	7.40	L	F	2002.02
851	22 May 2002	LTSB	North-West	Macclesfield, Cheshire	1300000	100000	7.69	L	F	2002.02
852	22 May 2002	RBS	South-East	Southend-on-Sea, Essex		164000				2002.02
853	24 Oct 2002	HBOS	South-East	Brighton, East Sussex	1300000	102787	7.91	L	F	2002.04
854	19 May 2003	LTSB	South-East	Headington, Oxon	1105000	84000	7.60	L	F	2003.02
855	23 Feb 2004	YB	North-East	Heckmondwike, W. Yorks	326000	18000	5.52	M	F	2004.01
856	23 Feb 2004	YB	North-East	Armley, Leeds, W. Yorks	365000	18000	4.93	M	F	2004.01
857	23 Feb 2004	YB	East Mid	Sutton-in-Ashfield, Nott	675000	37000	5.48	L	F	2004.01
858	17 May 2004	LTSB	South-West	Helston, Cornwall	370000	22000	5.95	M	F	2004.02
859	20 Oct 2004	HSBC	South-East	Bishop's Stortford, Hert	1155000	66500	5.76	L	F	2004.04
860	20 Oct 2004	YB	East Mid	Sutton-in-Ashfield, Nott	780000	37000	4.74	L	F	2004.04
861	17 Oct 2005	Barclay	London-M25	Plumstead SE18	624000	31500	5.05	L	F	2005.04
862	17 Oct 2005	Barclay	East Anglia	March, Cambs	765000	39100	5.11	L	F	2005.04
863	17 Oct 2005	Barclay	East Anglia	Swaffham, Norfolk	605000	32800	5.42	L	F	2005.04
864	17 Oct 2005	Barclay	East Anglia	Wisbech, Cambs	670000	32300	4.82	L	F	2005.04

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
865	17 Oct 2005	Barclay	East Anglia	Southwold, Suffolk	310000	12400	4.00	M	F	2005.04
866	17 Oct 2005	BB	South-East	Bexleyheath, Kent	877000	46000	5.25	L	F	2005.04
867	22 Feb 2006	Barclay	West Mid	Blackheath, Birmingham	1090000	50800	4.66	L	F	2006.01
868	22 Feb 2006	Barclay	East Mid	Kingsthorpe, Northants	640000	31500	4.92	L	F	2006.01
869	22 Feb 2006	Barclay	West Mid	Smethwick, Birmingham	730000	33900	4.64	L	F	2006.01
870	22 Feb 2006	Barclay	London-M25	Hayes, Middlesex	1515000	70000	4.62	L	F	2006.01
871	22 Feb 2006	Barclay	North-West	Prestwich, Manchester	560000	26000	4.64	L	F	2006.01
872	22 Feb 2006	Barclay	North-West	Macclesfield, Cheshire	840000	38000	4.52	L	F	2006.01
873	22 Feb 2006	Barclay	London-M25	Greenford, Middlesex	1170000	43000	3.68	L	F	2006.01
874	22 Feb 2006	Barclay	East Mid	Daventry, Northants	665000	29000	4.36	L	F	2006.01
875	22 Feb 2006	Barclay	South-East	Weybridge, Surrey	870000	37000	4.25	L	F	2006.01
876	22 Feb 2006	Barclay	East Mid	Melton Mowbray, Leics	680000	27800	4.09	L	F	2006.01
877	22 Feb 2006	Barclay	North-West	Blackburn, Lancashire	912000	41500	4.55	L	F	2006.01
878	22 Feb 2006	Barclay	North-West	Nantwich, Cheshire	800000	36000	4.50	L	F	2006.01
879	22 Feb 2006	Barclay	West Mid	Shirley, Solihull	1250000	57700	4.62	L	F	2006.01
880	22 Feb 2006	Barclay	South-East	Bracknell, Berks	1805000	100000	5.41	L	F	2006.01
881	22 Feb 2006	Barclay	South-East	Cranleigh, Surrey	940000	48000	5.11	L	F	2006.01
882	22 Feb 2006	Barclay	South-East	Godalming, Surrey	1105000	48000	4.34	L	F	2006.01
883	29 Mar 2006	Barclay	London-M25	Ealing SER, London	630000	27800	4.41	L	F	2006.01
884	29 Mar 2006	Barclay	East Anglia	Norwich AR, Norfolk	992000	57500	5.80	L	F	2006.01
885	29 Mar 2006	Barclay	South-West	Bridgwater, Somerset	990000	50000	5.05	L	F	2006.01
886	29 Mar 2006	Barclay	London-M25	Northwood, Middlesex	1206000	43000	3.57	L	F	2006.01
887	29 Mar 2006	Barclay	South-East	Royston, Hertfordshire	800000	35000	4.38	L	F	2006.01
888	29 Mar 2006	Barclay	North-East	Doncaster, S. Yorks	2150000	120000	5.58	L	F	2006.01
889	29 Mar 2006	Barclay	West Mid	Stone, Staffs	700000	33500	4.79	L	F	2006.01
890	29 Mar 2006	Barclay	South-East	Amersham, Bucks	901000	42500	4.72	L	F	2006.01
891	29 Mar 2006	Barclay	London-M25	Hornchurch, Essex	743000	37000	4.98	L	F	2006.01

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
892	29 Mar 2006	Barclay	East Anglia	Haverhill, Suffolk	945000	53500	5.66	L	F	2006.01
893	29 Mar 2006	Barclay	North-East	Scarborough, N.Yorks	1400000	70000	5.00	L	F	2006.01
894	29 Mar 2006	Barclay	North-West	Blackpool BS, Lancs	1350000	70000	5.19	L	F	2006.01
895	29 Mar 2006	Barclay	North-East	Scunthorpe, Lincs	1510000	85000	5.63	L	F	2006.01
896	29 Mar 2006	Barclay	London-M25	Chingford E4	855000	38200	4.47	L	F	2006.01
897	29 Mar 2006	Barclay	London-M25	South Woodford E18	666000	31400	4.71	L	F	2006.01
898	29 Mar 2006	Barclay	South-East	Sittingbourne, Kent	885000	38800	4.38	L	F	2006.01
899	29 Mar 2006	Barclay	South-East	Hitchin, Herts	1770000	70000	3.95	L	F	2006.01
900	29 Mar 2006	Barclay	North-West	Rochdale, Lancs	1180000	55000	4.66	L	F	2006.01
901	29 Mar 2006	Barclay	South-East	Southampton LR, Hants	1290000	68500	5.31	L	F	2006.01
902	29 Mar 2006	Barclay	East Mid	Grimsby, Lincs	1690000	95000	5.62	L	F	2006.01
903	29 Mar 2006	Barclay	South-East	Haywards Heath, W.Sussex	1650000	87500	5.30	L	F	2006.01
904	29 Mar 2006	Barclay	South-East	Stevenage, Herts	1735000	97500	5.62	L	F	2006.01
905	29 Mar 2006	Barclay	South-West	Falmouth, Cornwall	1120000	56500	5.04	L	F	2006.01
906	29 Mar 2006	Barclay	North-East	Wakefield, West Yorks	1500000	80000	5.33	L	F	2006.01
907	29 Mar 2006	Barclay	London-M25	Isleworth, Middlesex	820000	40700	4.96	L	F	2006.01
908	29 Mar 2006	Barclay	South-East	West Byfleet, Surrey	625000	34150	5.46	L	F	2006.01
909	29 Mar 2006	Barclay	South-East	Camberley, Surrey	1795000	88300	4.92	L	F	2006.01
910	29 Mar 2006	Barclay	North-East	Durham, Market Place	2950000	150000	5.08	L	F	2006.01
911	29 Mar 2006	Barclay	London-M25	Notting Hill, LadG	1850000	69000	3.73	L	F	2006.01
912	29 Mar 2006	Barclay	North-East	Leeds Park Sq, W. Yorks	1720000	69500	4.04	L	F	2006.01
913	29 Mar 2006	Barclay	South-West	Truro, Cornwall	2500000	125000	5.00	L	F	2006.01
914	29 Mar 2006	Barclay	London-M25	Whitechapel Rd, E1	2525000	100000	3.96	L	F	2006.01
915	29 Mar 2006	Barclay	South-East	Portsmouth North End	815000	41950	5.15	L	F	2006.01
916	29 Mar 2006	Barclay	South-West	St. Austell, Cornwall	850000	45000	5.29	L	F	2006.01
917	24 May 2006	Barclay	East Mid	West Bridgford, Notts	1225000	60940	4.97	L	L	2006.02
918	24 May 2006	Barclay	Wales	Merthyr Tydfil, W.Glam	1295000	66300	5.12	L	F	2006.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
919	24 May 2006	Barclay	Wales	Monmouth, Gwent	897000	47500	5.30	L	F	2006.02
920	24 May 2006	Barclay	East Mid	Northampton WR, Northant	1130000	55000	4.87	L	F	2006.02
921	24 May 2006	Barclay	East Mid	Dunstable, Beds	1265000	62500	4.94	L	F	2006.02
922	24 May 2006	Barclay	London-M25	Grays, Essex	917500	42000	4.58	L	F	2006.02
923	24 May 2006	Barclay	East Mid	Spalding, Lincs	1225000	67500	5.51	L	F	2006.02
924	24 May 2006	Barclay	East Anglia	Ely, Cambs	1255000	57500	4.60	L	F	2006.02
925	24 May 2006	Barclay	London-M25	Teddington, Middlesex	1200000	55400	4.60	L	F	2006.02
926	24 May 2006	Barclay	East Mid	Boston, Lincolnshire	1065000	52500	4.93	L	F	2006.02
927	24 May 2006	Barclay	London-M25	Stanmore, Middlesex	680000	28500	4.19	L	F	2006.02
928	24 May 2006	YB	North-East	Otley, West Yorks		33000			F	2006.02
929	24 May 2006	Barclay	South-West	Cirencester, Glos	1235000	60000	4.86	L	F	2006.02
930	24 May 2006	Barclay	London-M25	Golders Green, NW11	1155000	58000	5.02	L	F	2006.02
931	24 May 2006	Barclay	London-M25	Wimbledon	2245000	98000	4.37	L	F	2006.02
932	24 May 2006	Barclay	London-M25	Pinner, Middlesex	720000	35000	4.86	L	F	2006.02
933	24 May 2006	Barclay	South-East	Rickmansworth, Herts	1030000	45000	4.37	L	F	2006.02
934	24 May 2006	Barclay	South-East	Bognor Regis, W. Sussex	1020000	62000	6.08	L	F	2006.02
935	24 May 2006	Barclay	West Mid	Bromsgrove, Worcs	1460000	70000	4.79	L	F	2006.02
936	24 May 2006	Barclay	South-East	Andover, Hants	1520000	79800	5.25	L	F	2006.02
937	24 May 2006	Barclay	South-West	Weymouth, Dorset	1170000	60000	5.13	L	F	2006.02
938	24 May 2006	Woolwic	South-East	Southend-on-Sea, Essex	1795000	91000	5.16	L	F	2006.02
939	24 May 2006	Barclay	London-M25	Eltham, London	1765000	82700	4.69	L	F	2006.02
940	24 May 2006	Barclay	South-East	Bedford, Beds	1185000	60000	5.06	L	F	2006.02
941	24 May 2006	Barclay	London-M25	East Ham, London	1905000	81000	4.25	L	F	2006.02
942	24 May 2006	Barclay	South-East	Newbury, Berkshire	2280000	118000	5.18	L	F	2006.02
943	24 May 2006	Barclay	South-East	Barnet, Herts	1200000	48500	4.04	L	F	2006.02
944	06 Jul 2006	Barclay	Wales	Carmarthen, Dyfed	2130000	110000	5.16	L	F	2006.03
945	06 Jul 2006	Barclay	South-East	Harpenden, Herts	1900000	105000	5.53	L	L	2006.03

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
946	06 Jul 2006	Barclay	South-East	Epsom, Surrey	3000000	140000	4.67	L	F	2006.03
947	06 Jul 2006	Barclay	East Mid	Kettering, Northants	2525000	130000	5.15	L	F	2006.03
948	06 Jul 2006	Barclay	North-East	Bradford MP, W. Yorks	3360000	160000	4.76	L	F	2006.03
949	06 Jul 2006	Barclay	West Mid	Wolverhampton QS	2600000	125000	4.81	L	F	2006.03
950	06 Jul 2006	Barclay	South-East	Sutton, Surrey	2940000	130000	4.42	L	F	2006.03
951	06 Jul 2006	Barclay	East Mid	Worksop, Notts	1670000	90000	5.39	L	F	2006.03
952	06 Jul 2006	Barclay	South-East	Folkestone, Kent	1475000	70500	4.78	L	F	2006.03
953	06 Jul 2006	Barclay	London-M25	Barking, Essex	1600000	76000	4.75	L	F	2006.03
954	06 Jul 2006	BB	South-West	Swindon, Wiltshire		70000			F	2006.03
955	06 Jul 2006	Barclay	South-East	Brentwood, Essex	2775000	150000	5.41	L	F	2006.03
956	06 Jul 2006	Barclay	London-M25	Chadwell Heath, Essex	526000	22400	4.26	L	F	2006.03
957	06 Jul 2006	Barclay	North-East	Leeds H, West Yorks	2600000	105000	4.04	L	F	2006.03
958	06 Jul 2006	Barclay	West Mid	Coventry HS, W.Midlands	4000000	215000	5.38	L	F	2006.03
959	06 Jul 2006	Barclay	South-East	Billericay, Essex	1350000	71000	5.26	L	L	2006.03
960	06 Jul 2006	Barclay	South-West	Bournemouth CR, Dorset	3375000	150000	4.44	L	F	2006.03
961	06 Jul 2006	Barclay	North-West	Preston, Lancs	4450000	200000	4.49	L	F	2006.03
962	06 Jul 2006	Barclay	London-M25	Hampstead, London	3010000	155000	5.15	L	L	2006.03
963	06 Jul 2006	Barclay	South-East	High Wycombe, Bucks	2025000	105000	5.19	L	F	2006.03
964	06 Jul 2006	Barclay	London-M25	Uxbridge, Middlesex	3300000	175000	5.30	L	F	2006.03
965	06 Jul 2006	Barclay	South-East	Portsmouth, Hants	5310000	275000	5.18	L	F	2006.03
966	06 Jul 2006	Barclay	South-East	Witney, Oxfordshire	2430000	120000	4.94	L	F	2006.03
967	06 Jul 2006	Barclay	South-East	Wokingham, Berks	1575000	78000	4.95	L	F	2006.03
968	06 Jul 2006	Barclay	London-M25	Kew Gardens	1095000	55500	5.07	L	F	2006.03
969	23 May 2000	YB	North-East	Hunters Bar Sheffield		14000				2000.02
970	23 May 2000	YB	North-East	Gt. Horton, Bradford, WY		17000				2000.02
971	23 May 2000	YB	North-East	Queensbury, Bradford, WY		13000				2000.02
972	23 May 2000	YB	North-West	Shaw, Gt. Manchester		30000				2000.02

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
973	23 May 2000	YB	West Mid	Hanley, Stoke-on-Trent		64100				2000.02
974	23 May 2000	YB	North-East	Heckmondwike, W.Yorks		18000				2000.02
975	23 May 2000	YB	North-East	Malton, North Yorks		23000				2000.02
976	23 May 2000	YB	North-East	Chapel Allerton, Leeds		19000				2000.02
977	23 May 2000	YB	North-West	Cleverleys, Blackpool		11000				2000.02
978	23 May 2000	YB	North-West	Lytham St Annes, Lancs		29000				2000.02
979	23 May 2000	YB	North-East	Thornaby, Cleveland		21000				2000.02
980	23 May 2000	YB	North-West	Heywood, Manchester		14000				2000.02
981	23 May 2000	YB	North-East	Chapeltown, Sheffield		26000				2000.02
982	23 May 2000	YB	North-East	Armley, Leeds, W. Yorks		18000				2000.02
983	23 May 2000	YB	East Mid	Sutton-in-Ashfield, Nott		37000				2000.02
984	10 Jul 2000	YB	North-West	Staleybridge, Cheshire	150000	10000	6.67	S		2000.03
985	10 Jul 2000	YB	North-East	Thorpe, Doncaster, SY	160000	13000	8.13	S	F	2000.03
986	10 Jul 2000	YB	North-East	Mexborough, S.Yorks	177500	14000	7.89	S	F	2000.03
987	10 Jul 2000	YB	North-East	Goldthorpe, South Yorks	228000	18000	7.89	S	F	2000.03
988	10 Jul 2000	YB	North-West	Nelson, Lancs	250000	23000	9.20	S	L	2000.03
989	10 Jul 2000	YB	North-East	Batley, West Yorks	260000	17000	6.54	M	F	2000.03
990	10 Jul 2000	YB	North-West	Blackpool CD, Lancs	224000	15000	6.70	S	F	2000.03
991	10 Jul 2000	YB	North-East	Yealdon, Leeds, W. Yorks	264000	21000	7.95	M	F	2000.03
992	10 Jul 2000	YB	North-West	Littleborough, Lancashir	142000	16000	11.27	S	F	2000.03
993	10 Jul 2000	YB	North-East	Brighouse, Yorkshire	275000	20000	7.27	M	F	2000.03
994	10 Jul 2000	YB	North-East	Knaresborough, Yorks	240000	18000	7.50	S	L	2000.03
995	10 Jul 2000	YB	North-East	Cleckheaton, W. Yorks	242500	19000	7.84	S	F	2000.03
996	10 Jul 2000	YB	North-East	Guiseley, West Yorks	225000	14000	6.22	S	F	2000.03
997	27 Sep 2000	YB	North-West	Clitheroe, Lancs	322500	28000	8.68	M	F	2000.03
998	27 Sep 2000	YB	North-West	Accrington, Lancs	300000	30675	10.23	M	F	2000.03
999	27 Sep 2000	YB	North-East	Hillsborough, Sheffield	780000	70000	8.97	L	F	2000.03

Appendix III (Continued)

Original Full Dataset

Case No	Date	Bank	Region	Location	Price (£)	Rent (£)	Yield (%)	Lot Size	Tenure	Quarter
1000	27 Sep 2000	YB	North-West	Stockport PS, Cheshire	965000	82000	8.50	L	F	2000.03
1001	27 Sep 2000	YB	North-West	Wigan, Gt. Manchester	695000	57100	8.22	L	F	2000.03
1002	27 Sep 2000	YB	North-East	Whitby, North Yorks	470000	39560	8.42	M	F	2000.03
1003	27 Sep 2000	YB	North-East	Scarborough, N. Yorks	610000	57390	9.41	L	F	2000.03
1004	27 Sep 2000	YB	East Mid	Ripley, Derbyshire	310000	26000	8.39	M	F	2000.03
1005	27 Sep 2000	YB	East Mid	Long Eaton, Notts	335000	25000	7.46	M	F	2000.03
1006	27 Sep 2000	YB	North-East	Northallerton, N. Yorks	1160000	96750	8.34	L	F	2000.03
1007	27 Sep 2000	YB	North-East	Skipton, North Yorks	675000	59400	8.80	L	F	2000.03
1008	27 Sep 2000	YB	North-East	Hartlepool, Clevedon	685000	59500	8.69	L	F	2000.03
1009	27 Sep 2000	YB	North-East	Beverley, Yorks	750000	65000	8.67	L	F	2000.03
1010	27 Sep 2000	YB	West Mid	Leamington Spa, Warks	490000	38000	7.76	M	F	2000.03
1011	27 Sep 2000	YB	North-East	Bishop Auckland Cty Durh	410000	37000	9.02	M	F	2000.03
1012	27 Sep 2000	YB	North-East	Thirsk, North Yorkshire	295000	22750	7.71	M	F	2000.03

Appendix IV

Histogram of Yield – Three Main Banks datasets at increments of 0.25 per cent

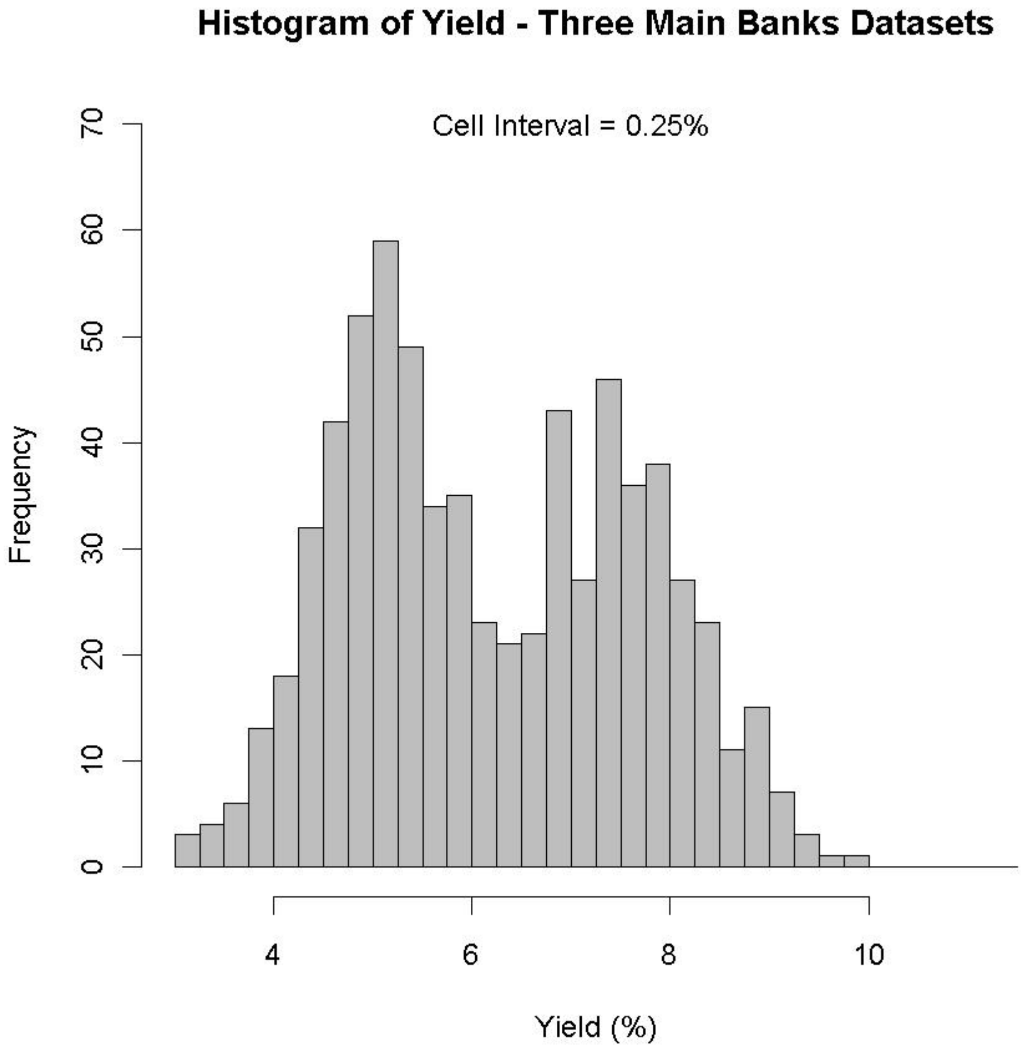


Figure IV.1 Histogram of Yield – Three Main Banks datasets at increments of 0.25 per cent

The histogram in Figure IV.1 is exactly the same shape as the one in Figure 8.6, because frequency and probability density are linearly related.

Appendix V

Crosstabulations

Table V.1 Crosstabulation of the *Three Main Banks C* dataset (Provinces B)

	Yield	Bank	Time	Lot Size	Provinces B	Frequency
1	Low	Barclays	Early	Small	London & SE	0
2	High	Barclays	Early	Small	London & SE	0
3	Low	HSBC	Early	Small	London & SE	3
4	High	HSBC	Early	Small	London & SE	26
5	Low	Lloyds TSB	Early	Small	London & SE	17
6	High	Lloyds TSB	Early	Small	London & SE	29
7	Low	Barclays	Late	Small	London & SE	11
8	High	Barclays	Late	Small	London & SE	0
9	Low	HSBC	Late	Small	London & SE	3
10	High	HSBC	Late	Small	London & SE	0
11	Low	Lloyds TSB	Late	Small	London & SE	13
12	High	Lloyds TSB	Late	Small	London & SE	0
13	Low	Barclays	Early	Large	London & SE	0
14	High	Barclays	Early	Large	London & SE	1
15	Low	HSBC	Early	Large	London & SE	0
16	High	HSBC	Early	Large	London & SE	2
17	Low	Lloyds TSB	Early	Large	London & SE	10
18	High	Lloyds TSB	Early	Large	London & SE	17
19	Low	Barclays	Late	Large	London & SE	85
20	High	Barclays	Late	Large	London & SE	1
21	Low	HSBC	Late	Large	London & SE	10
22	High	HSBC	Late	Large	London & SE	2
23	Low	Lloyds TSB	Late	Large	London & SE	11
24	High	Lloyds TSB	Late	Large	London & SE	1
25	Low	Barclays	Early	Small	Wales & SW	0
26	High	Barclays	Early	Small	Wales & SW	0
27	Low	HSBC	Early	Small	Wales & SW	3
28	High	HSBC	Early	Small	Wales & SW	47
29	Low	Lloyds TSB	Early	Small	Wales & SW	25
30	High	Lloyds TSB	Early	Small	Wales & SW	40
31	Low	Barclays	Late	Small	Wales & SW	2
32	High	Barclays	Late	Small	Wales & SW	0
33	Low	HSBC	Late	Small	Wales & SW	4
34	High	HSBC	Late	Small	Wales & SW	0
35	Low	Lloyds TSB	Late	Small	Wales & SW	8
36	High	Lloyds TSB	Late	Small	Wales & SW	0
37	Low	Barclays	Early	Large	Wales & SW	0
38	High	Barclays	Early	Large	Wales & SW	1
39	Low	HSBC	Early	Large	Wales & SW	0

40	High	HSBC	Early	Large	Wales & SW	1
41	Low	Lloyds TSB	Early	Large	Wales & SW	2
42	High	Lloyds TSB	Early	Large	Wales & SW	10
43	Low	Barclays	Late	Large	Wales & SW	17
44	High	Barclays	Late	Large	Wales & SW	0
45	Low	HSBC	Late	Large	Wales & SW	4
46	High	HSBC	Late	Large	Wales & SW	0
47	Low	Lloyds TSB	Late	Large	Wales & SW	4
48	High	Lloyds TSB	Late	Large	Wales & SW	0
49	Low	Barclays	Early	Small	Midlands	0
50	High	Barclays	Early	Small	Midlands	0
51	Low	HSBC	Early	Small	Midlands	0
52	High	HSBC	Early	Small	Midlands	9
53	Low	Lloyds TSB	Early	Small	Midlands	17
54	High	Lloyds TSB	Early	Small	Midlands	37
55	Low	Barclays	Late	Small	Midlands	3
56	High	Barclays	Late	Small	Midlands	0
57	Low	HSBC	Late	Small	Midlands	1
58	High	HSBC	Late	Small	Midlands	0
59	Low	Lloyds TSB	Late	Small	Midlands	8
60	High	Lloyds TSB	Late	Small	Midlands	0
61	Low	Barclays	Early	Large	Midlands	0
62	High	Barclays	Early	Large	Midlands	0
63	Low	HSBC	Early	Large	Midlands	0
64	High	HSBC	Early	Large	Midlands	0
65	Low	Lloyds TSB	Early	Large	Midlands	3
66	High	Lloyds TSB	Early	Large	Midlands	15
67	Low	Barclays	Late	Large	Midlands	39
68	High	Barclays	Late	Large	Midlands	0
69	Low	HSBC	Late	Large	Midlands	10
70	High	HSBC	Late	Large	Midlands	0
71	Low	Lloyds TSB	Late	Large	Midlands	7
72	High	Lloyds TSB	Late	Large	Midlands	3
73	Low	Barclays	Early	Small	North Britain	0
74	High	Barclays	Early	Small	North Britain	0
75	Low	HSBC	Early	Small	North Britain	0
76	High	HSBC	Early	Small	North Britain	24
77	Low	Lloyds TSB	Early	Small	North Britain	4
78	High	Lloyds TSB	Early	Small	North Britain	28
79	Low	Barclays	Late	Small	North Britain	0
80	High	Barclays	Late	Small	North Britain	0
81	Low	HSBC	Late	Small	North Britain	2
82	High	HSBC	Late	Small	North Britain	0
83	Low	Lloyds TSB	Late	Small	North Britain	3
84	High	Lloyds TSB	Late	Small	North Britain	3
85	Low	Barclays	Early	Large	North Britain	0
86	High	Barclays	Early	Large	North Britain	2
87	Low	HSBC	Early	Large	North Britain	1
88	High	HSBC	Early	Large	North Britain	0
89	Low	Lloyds TSB	Early	Large	North Britain	0

90	High	Lloyds TSB	Early	Large	North Britain	17
91	Low	Barclays	Late	Large	North Britain	30
92	High	Barclays	Late	Large	North Britain	0
93	Low	HSBC	Late	Large	North Britain	8
94	High	HSBC	Late	Large	North Britain	0
95	Low	Lloyds TSB	Late	Large	North Britain	4
96	High	Lloyds TSB	Late	Large	North Britain	3

Table V.2 Crosstabulation of *Region* and *Provinces A*

Region * Provinces A Crosstabulation

Count

		Provinces A				
		North	Midlands	South	Celt	Total
Region	London-M25	0	0	48	0	48
	South-East	0	0	194	0	194
	South-West	0	0	103	0	103
	East Anglia	0	32	0	0	32
	West Midlands	0	67	0	0	67
	East Midlands	0	53	0	0	53
	North-East	61	0	0	0	61
	North-West	53	0	0	0	53
	Wales	0	0	0	65	65
	Scotland	0	0	0	15	15
	Total	114	152	345	80	691

Table V.3 Crosstabulation of *Region* and *Provinces B*

Region * Provinces B Crosstabulation

Count

		Provinces B				
		London & SE	Wales & SW	Midlands	North Britain	Total
Region	London-M25	48	0	0	0	48
	South-East	194	0	0	0	194
	South-West	0	103	0	0	103
	East Anglia	0	0	32	0	32
	West Midlands	0	0	67	0	67
	East Midlands	0	0	53	0	53
	North-East	0	0	0	61	61
	North-West	0	0	0	53	53
	Wales	0	65	0	0	65
	Scotland	0	0	0	15	15
	Total	242	168	152	129	691

Appendix VI

The Shapiro-Wilk normality tests output

Figure VI.1 Shapiro-Wilk normality test with respect to *Provinces A*

```
        Shapiro-Wilk normality test Provinces A


 yields<-testframe$yieldvals[testframe$group==1]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9795, p-value = 0.8868


 yields<-testframe$yieldvals[testframe$group==2]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9577, p-value = 0.2701



 yields<-testframe$yieldvals[testframe$group==3]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.911, p-value = 0.006056

 qqnorm(yields)

 yields<-testframe$yieldvals[testframe$group==4]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9675, p-value = 0.313


 yields<-testframe$yieldvals[testframe$group==5]
 out<-shapiro.test(yields)
```

```
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9727, p-value = 0.3335


 yields<-testframe$yieldvals[testframe$group==6]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9637, p-value = 0.224


 yields<-testframe$yieldvals[testframe$group==7]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.933, p-value = 0.00973


 yields<-testframe$yieldvals[testframe$group==8]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9704, p-value = 0.678


 yields<-testframe$yieldvals[testframe$group==9]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9839, p-value = 0.2762


 yields<-testframe$yieldvals[testframe$group==10]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9629, p-value = 0.4511
```

```
 yields<-testframe$yieldvals[testframe$group==11]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9415, p-value = 0.09062
```

Figure VI.2 Shapiro-Wilk normality test with respect to *Provinces B*

```
        Shapiro-Wilk normality test Provinces B


 #Test of Normality

 yields<-testframe$yieldvals[testframe$group==1]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9795, p-value = 0.8868


 yields<-testframe$yieldvals[testframe$group==2]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9577, p-value = 0.2701


 yields<-testframe$yieldvals[testframe$group==3]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.911, p-value = 0.006056


 yields<-testframe$yieldvals[testframe$group==4]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9675, p-value = 0.313


 yields<-testframe$yieldvals[testframe$group==5]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9727, p-value = 0.6938
```

```
 yields<-testframe$yieldvals[testframe$group==6]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.935, p-value = 0.1132


 yields<-testframe$yieldvals[testframe$group==7]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9366, p-value = 0.08191


 yields<-testframe$yieldvals[testframe$group==8]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9354, p-value = 0.0842


 yields<-testframe$yieldvals[testframe$group==9]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9856, p-value = 0.4685


 yields<-testframe$yieldvals[testframe$group==10]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test

data:  yields
W = 0.9758, p-value = 0.4317


 yields<-testframe$yieldvals[testframe$group==11]
 out<-shapiro.test(yields)
 out

        Shapiro-Wilk normality test
```

```
data:  yields
W = 0.9254, p-value = 0.01141
```

Appendix VII

Provinces A logistic regression output

```
#
> # Malvern Tipping logistic cross validation Analysis in R
> #
>
>
>
#**********************************************************************
*************************
> #
> #            remove "#"s to output text to file
> #
> #    NB To print out figures must still step through script using
CTRL-R so have time to
> #    access "file" header to output graphs to file
> #
>
#**********************************************************************
*************************
> #sink(file="logisticcrossvalidationProvA.txt")
>
>
>
#**********************************************************************
*************************
> #
> # Load package 'foreign' to access read.spss function
> #
>
#**********************************************************************
*************************
> library(foreign)
>
>
>
#**********************************************************************
*************************
> #
> #         read data from an SPSS file and convert to a data.frame
> #
>
#**********************************************************************
*************************
> filename<-"Three Main Banks B.sav"
> RawData<-read.spss(filename,to.data.frame=TRUE)
>
> writeLines("\n\n\nNames of variables in Input File\n\n\n")



Names of variables in Input File



> names(RawData)
 [1] "CaseNo"     "Date"       "Bank"       "Region"     "Location"
 [6] "Price"      "Rent"       "Yield"      "Lotsize"    "Tenure"
[11] "Quarter"    "BankNo"     "RegCode"    "SizeCode"   "TenuCode"
[16] "Year"       "ProvincesA" "YieldGroup" "LotGroup"   "Time"
>
> writeLines("\n\n\nSummary of Input Data\n\n\n")
```

Summary of Input Data

```
> summary(RawData,maxsum=25)
     CaseNo                 Date             Bank              Region   
 Min.   :  1.0   15 Jul 1997: 97   Barclay:192   East Anglia: 32  
 1st Qu.:254.5   27 Jun 2001: 92   HSBC   :160   East Mid   : 53  
 Median :520.0   21 Nov 2001: 54   LTSB   :339   London-M25 : 49  
 Mean   :504.1   29 Mar 2006: 34                 North-East : 61  
 3rd Qu.:762.5   05 Jul 2006: 25                 North-West : 53  
 Max.   :968.0   24 May 2006: 24                 Scotland   : 15  
                 16 May 2006: 23                 South-East :193  
                 06 Jul 2006: 21                 South-West :103  
                 06 Jul 2005: 20                 Wales      : 65  
                 03 Jul 2006: 19                 West Mid   : 67  
                 28 Mar 2006: 19                                  
                 17 Oct 2001: 17                                  
                 22 Feb 2006: 16                                  
                 19 Mar 2002: 13                                  
                 22 Mar 2006: 13                                  
                 08 Oct 2003: 12                                  
                 10 Feb 2004: 12                                  
                 19 May 2004: 11                                  
                 22 May 2001: 11                                  
                 25 May 2006: 11                                  
                 23 May 2002: 10                                  
                 25 Mar 2004: 10                                  
                 04 Feb 2002:  9                                  
                 03 Dec 2003:  8                                  
                 (Other)    :110                                  
                     Location       Price               Rent       
 Chatham, Kent            :  4   Min.   :  69000   Min.   :  4400  
 Melton Mowbray, Leics    :  4   1st Qu.: 250000   1st Qu.: 16500  
 Tadley, Hampshire        :  4   Median : 463000   Median : 27000  
 Brixham, Devon           :  3   Mean   : 650092   Mean   : 37262  
 Coulsdon, Surrey         :  3   3rd Qu.: 810000   3rd Qu.: 45000  
 Crediton, Devon          :  3   Max.   :5310000   Max.   :275000  
 Cullumpton, Devon        :  3                                     
 East Ham, London         :  3                                     
 Maldon, Essex            :  3                                     
 Okehampton, Devon        :  3                                     
 Wareham, Dorset          :  3                                     
 Amesbury, Wiltshire      :  2                                     
 Ammanford, Dyfed         :  2                                     
 Aylsham, Norfolk         :  2                                     
 Bala, Gwynedd            :  2                                     
 Bedford, Beds            :  2                                     
 Bedminster, Avon         :  2                                     
 Bicester, Oxfordshire    :  2                                     
 Bodmin, Cornwall         :  2                                     
 Boscombe, Bournemouth    :  2                                     
 Bridgwater, Somerset     :  2                                     
 Bromyard, Herefordshire :  2                                     
 Camberley, Surrey        :  2                                     
 Cleckheaton, West Yorks :  2                                     
 (Other)                  :629                                     
     Yield           Lotsize         Tenure         Quarter    
 Min.   :3.170   L       :316   F       :679   Min.   :1997  
 1st Qu.:5.040   M       :201   Feudal  :  2   1st Qu.:2001  
 Median :6.000   S       :174   Feuhold : 10   Median :2002  
 Mean   :6.223                                 Mean   :2003  
 3rd Qu.:7.460                                 3rd Qu.:2006  
 Max.   :9.900                                 Max.   :2006  
```

```
                    BankNo             RegCode       SizeCode
TenuCode
 hsbc                 :160   London M25   : 48   small :173   F
:679
 lloydstsb            :339   South east   :194   medium:197   feuhold
: 12
 natwest              :  0   South west   :103   large :321   L
:  0
 rbs                  :  0   East anglia  : 32
Heritable:  0
 BW                   :  0   West midlands: 67
 BB                   :  0   East Midlands: 53
 barclays             :192   North east   : 61
 yorkshire            :  0   North west   : 53
 woolwich             :  0   Wales        : 65
 cheltenham and glos  :  0   Scotland     : 15
 abbey national       :  0
 clydesdale           :  0
 halifax/bos          :  0
 alliance and leicester:  0
```

```
     Year           ProvincesA  YieldGroup  LotGroup      Time
 Min.   :1997   North   :114   Low :372   Small:370   Early:391
 1st Qu.:2001   Midlands:152   High:319   Large:321   Late :300
 Median :2002   South   :345
 Mean   :2003   Celt    : 80
 3rd Qu.:2006
 Max.   :2006
```

```
>
>
#*********************************************************************
************************
> #
> #     Extract required data into Operating Data File
> #
>
#*********************************************************************
************************
>
>
> opdata<-with(RawData,{
+ bank<-factor(Bank)
+ time<-factor(Time)
+ provincesa<-factor(ProvincesA)
+ yield<-factor(YieldGroup)
+ lotsize<-factor(LotGroup)
+ oplist<-list(yield=as.factor(yield),bank=as.factor(bank),
+  time=as.factor(time),lotsize=as.factor(lotsize),
+  provincesa=as.factor(provincesa))
+ opdata<-as.data.frame(oplist)
+ return(opdata)})
>
> writeLines("\n\n\nSummary of Operating Data used in Analysis\n\n\n")



Summary of Operating Data used in Analysis



> summary(opdata)
  yield         bank        time      lotsize       provincesa
 Low :372   Barclay:192   Early:391   Small:370   North   :114
 High:319   HSBC   :160   Late :300   Large:321   Midlands:152
            LTSB   :339                           South   :345
                                                  Celt    : 80
>
>
>
>
#*********************************************************************
************************
> #
> #     Basic Model using step for backwards elimination
> #
>
#*********************************************************************
************************
>
> modl0<-
glm(yield~(bank+time+lotsize+provincesa)^4,family=binomial,opdata)
Warning messages:
1: In glm.fit(x = X, y = Y, weights = weights, start = start, etastart
= etastart,  :
  algorithm did not converge
2: In glm.fit(x = X, y = Y, weights = weights, start = start, etastart
= etastart,  :
  fitted probabilities numerically 0 or 1 occurred
>
> writeLines("\n\n\nBasic Model after backwards elimination from
Saturated Model\nusing Akaike's Information Criterion\n\n\n")
```

Basic Model after backwards elimination from Saturated Model
using Akaike's Information Criterion

```
> modl1<-step(modl0)
Start:  AIC=468.59
yield ~ (bank + time + lotsize + provincesa)^4

                               Df Deviance    AIC
- bank:time:lotsize:provincesa  3   383.37 459.37
<none>                              386.59 468.59

Step:  AIC=459.37
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize +
    bank:provincesa + time:lotsize + time:provincesa +
lotsize:provincesa +
    bank:time:lotsize + bank:time:provincesa + bank:lotsize:provincesa
+
    time:lotsize:provincesa

                          Df Deviance    AIC
- bank:time:provincesa     4   383.37 451.37
- bank:time:lotsize        1   383.37 457.37
- time:lotsize:provincesa  3   388.98 458.98
- bank:lotsize:provincesa  5   393.03 459.03
<none>                         383.37 459.37

Step:  AIC=451.37
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize +
    bank:provincesa + time:lotsize + time:provincesa +
lotsize:provincesa +
    bank:time:lotsize + bank:lotsize:provincesa +
time:lotsize:provincesa

                          Df Deviance    AIC
- bank:time:lotsize        1   383.37 449.37
- time:lotsize:provincesa  3   388.98 450.98
<none>                         383.37 451.37
- bank:lotsize:provincesa  5   402.30 460.30

Step:  AIC=449.37
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize +
    bank:provincesa + time:lotsize + time:provincesa +
lotsize:provincesa +
    bank:lotsize:provincesa + time:lotsize:provincesa

                          Df Deviance    AIC
- time:lotsize:provincesa  3   388.98 448.98
<none>                         383.37 449.37
- bank:time                2   394.93 456.93
- bank:lotsize:provincesa  5   403.97 459.97

Step:  AIC=448.98
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize +
    bank:provincesa + time:lotsize + time:provincesa +
lotsize:provincesa +
    bank:lotsize:provincesa

                          Df Deviance    AIC
- time:provincesa          3   389.76 443.76
- time:lotsize             1   389.09 447.09
<none>                         388.98 448.98
- bank:lotsize:provincesa  5   407.87 457.87
- bank:time                2   405.08 461.08

Step:  AIC=443.76
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize +
```

```
    bank:provincesa + time:lotsize + lotsize:provincesa +
bank:lotsize:provincesa

                          Df Deviance    AIC
- time:lotsize             1   389.83 441.83
<none>                         389.76 443.76
- bank:lotsize:provincesa  5   408.21 452.21
- bank:time                2   407.31 457.31

Step:  AIC=441.83
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize +
    bank:provincesa + lotsize:provincesa + bank:lotsize:provincesa

                          Df Deviance    AIC
<none>                         389.83 441.83
- bank:lotsize:provincesa  5   409.85 451.85
- bank:time                2   407.39 455.39
There were 23 warnings (use warnings() to see them)
>
>
>
>
#***********************************************************************
************************
> #
> #     Confusion Matrix for Basic Model p=0.5
> #
>
#***********************************************************************
************************
> writeLines("\n\n\nConfusion Matrix for Basic Model p=0.5\n\n\n")



Confusion Matrix for Basic Model p=0.5



>
> pred<-(rep("Low",length(fitted(modl1))))
> pred<-factor(pred,levels=c("Low","High"))
> pred[fitted(modl1)<0.5]<-"Low"
> pred[fitted(modl1)>=0.5]<-"High"
> table(pred,yield=opdata$yield)
      yield
pred   Low High
  Low  288   12
  High  84  307
>
>
>
#***********************************************************************
************************
> #
> #     Co-efficients for fitted Basic Model
> #
>
#***********************************************************************
************************
> writeLines("\n\n\nCo-efficients for Basic Model\n\n\n")



Co-efficients for Basic Model



> coef(modl1)
                                (Intercept)
```

```
                              35.6205854
                             bankHSBC
                             -16.4378680
                             bankLTSB
                             -33.7082204
                                timeLate
                             -38.1684483
                            lotsizeLarge
                             -16.9344411
                       provincesaMidlands
                             -17.0182056
                          provincesaSouth
                             -17.0182056
                           provincesaCelt
                             -17.0182056
                    bankHSBC   :timeLate
                               1.4481164
                    bankLTSB   :timeLate
                              34.8196017
                bankHSBC   :lotsizeLarge
                             -19.1705793
                bankLTSB   :lotsizeLarge
                              18.0882619
          bankHSBC   :provincesaMidlands
                              16.9447584
          bankLTSB   :provincesaMidlands
                              15.8372194
             bankHSBC   :provincesaSouth
                              -0.1061236
             bankLTSB   :provincesaSouth
                              15.2334117
              bankHSBC   :provincesaCelt
                              17.1765388
              bankLTSB   :provincesaCelt
                              17.8243755
         lotsizeLarge:provincesaMidlands
                              16.9344411
            lotsizeLarge:provincesaSouth
                              31.9155422
             lotsizeLarge:provincesaCelt
                              16.9344411
bankHSBC   :lotsizeLarge:provincesaMidlands
                              17.2155725
bankLTSB   :lotsizeLarge:provincesaMidlands
                             -16.7954764
   bankHSBC   :lotsizeLarge:provincesaSouth
                              37.1466739
   bankLTSB   :lotsizeLarge:provincesaSouth
                             -32.5093751
    bankHSBC   :lotsizeLarge:provincesaCelt
                              17.7839076
    bankLTSB   :lotsizeLarge:provincesaCelt
                                      NA
> anova(modl1,test="Chisq")
Analysis of Deviance Table

Model: binomial, link: logit

Response: yield

Terms added sequentially (first to last)


                        Df Deviance Resid. Df Resid. Dev P(>|Chi|)
NULL                                      690     953.86
bank                     2  253.745       688     700.12 < 2.2e-16 ***
time                     1  211.606       687     488.51 < 2.2e-16 ***
lotsize                  1    3.285       686     485.22  0.069906 .
provincesa               3   39.545       683     445.68 1.330e-08 ***
```

```
bank:time               2   28.771       681     416.91 5.657e-07 ***
bank:lotsize            2    1.079       679     415.83  0.582953
bank:provincesa         6    5.201       673     410.63  0.518348
lotsize:provincesa      3    0.774       670     409.85  0.855669
bank:lotsize:provincesa 5   20.022       665     389.83  0.001238 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Warning messages:
1: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
2: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
> summary(modl1)

Call:
glm(formula = yield ~ bank + time + lotsize + provincesa + bank:time +
    bank:lotsize + bank:provincesa + lotsize:provincesa +
bank:lotsize:provincesa,
    family = binomial, data = opdata)

Deviance Residuals:
       Min          1Q      Median          3Q         Max
-2.359e+00  -1.425e-01  -7.976e-05   4.902e-01   3.032e+00

Coefficients: (1 not defined because of singularities)
                                              Estimate Std. Error   z value
(Intercept)                                    35.6206 13409.7931     0.003
bankHSBC                                      -16.4379 13510.5094    -0.001
bankLTSB                                      -33.7082 13409.7931    -0.003
timeLate                                      -38.1684  5247.0558    -0.007
lotsizeLarge                                  -16.9344 12483.4356    -0.001
provincesaMidlands                            -17.0182 14068.7941    -0.001
provincesaSouth                               -17.0182 13000.7218    -0.001
provincesaCelt                                -17.0182  6613.0606    -0.003
bankHSBC   :timeLate                            1.4481  5667.6181  2.56e-04
bankLTSB   :timeLate                           34.8196  5247.0558     0.007
bankHSBC   :lotsizeLarge                      -19.1706 12913.9190    -0.001
bankLTSB   :lotsizeLarge                       18.0883 12483.4355     0.001
bankHSBC   :provincesaMidlands                 16.9448 14341.0231     0.001
bankLTSB   :provincesaMidlands                 15.8372 14068.7941     0.001
bankHSBC   :provincesaSouth                    -0.1061 13104.5824 -8.10e-06
bankLTSB   :provincesaSouth                    15.2334 13000.7218     0.001
bankHSBC   :provincesaCelt                     17.1765  6989.2690     0.002
bankLTSB   :provincesaCelt                     17.8244  6613.0605     0.003
lotsizeLarge:provincesaMidlands                16.9344 14048.1712     0.001
lotsizeLarge:provincesaSouth                   31.9155 12863.6528     0.002
lotsizeLarge:provincesaCelt                    16.9344  3357.8515     0.005
bankHSBC   :lotsizeLarge:provincesaMidlands    17.2156 15138.7751     0.001
bankLTSB   :lotsizeLarge:provincesaMidlands   -16.7955 14048.1712    -0.001
bankHSBC   :lotsizeLarge:provincesaSouth       37.1467 13556.0191     0.003
bankLTSB   :lotsizeLarge:provincesaSouth      -32.5094 12863.6528    -0.003
bankHSBC   :lotsizeLarge:provincesaCelt        17.7839  6038.3476     0.003
bankLTSB   :lotsizeLarge:provincesaCelt             NA         NA        NA
                                            Pr(>|z|)
(Intercept)                                    0.998
bankHSBC                                       0.999
bankLTSB                                       0.998
timeLate                                       0.994
lotsizeLarge                                   0.999
provincesaMidlands                             0.999
provincesaSouth                                0.999
provincesaCelt                                 0.998
bankHSBC   :timeLate                           1.000
bankLTSB   :timeLate                           0.995
bankHSBC   :lotsizeLarge                       0.999
bankLTSB   :lotsizeLarge                       0.999
bankHSBC   :provincesaMidlands                 0.999
bankLTSB   :provincesaMidlands                 0.999
bankHSBC   :provincesaSouth                    1.000
bankLTSB   :provincesaSouth                    0.999
bankHSBC   :provincesaCelt                     0.998
bankLTSB   :provincesaCelt                     0.998
lotsizeLarge:provincesaMidlands                0.999
lotsizeLarge:provincesaSouth                   0.998
lotsizeLarge:provincesaCelt                    0.996
bankHSBC   :lotsizeLarge:provincesaMidlands    0.999
bankLTSB   :lotsizeLarge:provincesaMidlands    0.999
bankHSBC   :lotsizeLarge:provincesaSouth       0.998
```

```
bankLTSB   :lotsizeLarge:provincesaSouth       0.998
bankHSBC   :lotsizeLarge:provincesaCelt        0.998
bankLTSB   :lotsizeLarge:provincesaCelt           NA

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 953.86  on 690  degrees of freedom
Residual deviance: 389.83  on 665  degrees of freedom
AIC: 441.83

Number of Fisher Scoring iterations: 18

>
>
>
>
>
#*****************************************************************************************************
> #
> #     Cross Validation of Basic Model with data divided randomly into
10 approx.equal groups
> #
>
#*****************************************************************************************************
>
>
>
>
>
#*****************************************************************************************************
> #     Randomly divide data into 10 equal groups
>
#*****************************************************************************************************
>
> randnos<-with(opdata,runif(length(yield)))
> orderednos<-order(randnos)
> full.length<-with(opdata,full.length<-(length(yield)))
> sub.length<-floor(full.length/10)
> sections<-seq(0,full.length,sub.length)
> sections[11]<-full.length
>
>
>
#*****************************************************************************************************
> #     Cross Validate
>
#*****************************************************************************************************
>
> model.formula<-formula(modl1)
> cross.val.data<-cbind(opdata,randnos,orderednos)
>
> for(i in 1:10){
+ cross.val.sample<-with(cross.val.data,{
+ tempdata0<-cross.val.data[{orderednos<=sections[i] |
orderednos>sections[i+1]},]
+ return(tempdata0)
+ })
+
+ cross.val.test<-with(cross.val.data,{
+ tempdata1<-cross.val.data[{orderednos>sections[i] &
orderednos<=sections[i+1]},]
+ return(tempdata1)
+ })
+
+
```

```
+ fitted.model<-glm(model.formula,family=binomial,cross.val.sample)
+
+ fitted.values<-
predict(fitted.model,newdata=cross.val.test,type="response")
+ outfile<-cbind(cross.val.test,fitted=fitted.values)
+
+ if(i==1)results<-outfile
+ else results<-rbind(results,outfile)
+
+
+ }
There were 20 warnings (use warnings() to see them)
>
>
#***********************************************************************
************************
> #
> #     Confusion Matrix for Basic Model after Cross Validation
> #
>
#***********************************************************************
************************
> writeLines("\n\n\nConfusion Matrix after Cross Validation - Basic
Model p=0.5\n\n\n")



Confusion Matrix after Cross Validation - Basic Model p=0.5



>
> pred<-(rep("Low",length(results$fitted)))
> pred<-factor(pred,levels=c("Low","High"))
> pred[results$fitted<0.5]<-"Low"
> pred[results$fitted>=0.5]<-"High"
>
> table(pred,yield=results$yield)
      yield
pred   Low High
  Low  286   16
  High  86  303
>
>
>
#***********************************************************************
************************
> #
> #     ROC chart for Basic Model after Cross Validation
> #
>
#***********************************************************************
************************
> writeLines("\n\n\nROC Chart after Cross Validation - Basic
Model\n\n\n")



ROC Chart after Cross Validation - Basic Model



>
> library(ROCR)
> basic.pred<-
prediction(results$fitted,results$yield,label.ordering=c("Low","High"))
> perf<-performance(basic.pred,"tpr","fpr")
> plot(perf,main="ROC Chart after Cross Validation - Basic Model
Provinces A",cex.main=0.9,colorize=TRUE)
```

```
> text(locator(1),"Key to Cutoff Probability",srt=90)
>
>
>
#*********************************************************************
************************
> #     Output ROC chart data for Basic Model after Cross Validation
>
#*********************************************************************
************************
> rocout<-data.frame(fitted=results$fitted,yield=results$yield)
> write.table(rocout,file="provabasicroc.out")
>
> #NB ROC data is also sufficient to reproduce the bar chart.
>
>
#*********************************************************************
*******
> #
> #                                bar chart
> #
>
#*********************************************************************
*******
>
> limits<-seq(0.0,1.0,0.1)
> temppred<-rep(1,length(results$fitted))
>
> predgrp<-with(results,{
+ for(i in 1:10)temppred[fitted>limits[i] & fitted<=limits[i+1]]<-i
+ return(temppred)
+ }
+ )
>
> dummy<-rep(1,length(predgrp))
> add.results<-cbind(results,pred.grp=predgrp,counter=dummy)
>
> #"data" is the array containing the number of low yield and high
yield observations
> # whose prdicted probabilites of being in thehigh yield group are in
the ranges
> #specified by "intervals"
>
> barchart.data<-
with(add.results,tapply(counter,list(yield,pred.grp),sum))
>
> intervals<-c("0.00-0.09","0.10-0.19","0.20-0.29","0.30-0.39","0.40-
0.49",
+  "0.50-0.59","0.60-0.69","0.70-0.79","0.80-0.89","0.90-1.00")
>
>
barplot(barchart.data,beside=TRUE,names.arg=intervals,space=c(1,5),cex.
names=0.5,
+   xlab="Probability case belongs to high yield group",
+   ylab="No of cases",
+   col=c("grey25","grey75"),
+   main="Counts against Probability(high yield) Provinces
A",cex.main=0.9,
+         sub="Basic Model",cex.sub=0.9)
>
> legend(locator(1),legend=c("low yield","high
yield"),fill=c("grey25","grey75"))
>
>
>
>
>
#*********************************************************************
*********
```

```
> #
> #    Simplified Model excluding higher order term
> #
>
#*********************************************************************
*********
>
> writeLines("\n\n\nExamine STATISTICAL significance of third order
term bank:lotsize:provinces\n\n\n")



Examine STATISTICAL significance of third order term
bank:lotsize:provinces



> modl2<-update(modl1,~.-bank:lotsize:provincesa)
Warning message:
In glm.fit(x = X, y = Y, weights = weights, start = start, etastart =
etastart,  :
  fitted probabilities numerically 0 or 1 occurred
> anova(modl1,modl2,test="Chi")
Analysis of Deviance Table

Model 1: yield ~ bank + time + lotsize + provincesa + bank:time +
bank:lotsize +
    bank:provincesa + lotsize:provincesa + bank:lotsize:provincesa
Model 2: yield ~ bank + time + lotsize + provincesa + bank:time +
bank:lotsize +
    bank:provincesa + lotsize:provincesa
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       665     389.83
2       670     409.85 -5  -20.022  0.001238 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> writeLines("\n\n\nAlthough term is STATISTICALLY significant remove
it and see if it has PRACTICAL significance\n\n\n")



Although term is STATISTICALLY significant remove it and see if it has
PRACTICAL significance



> modl3<-step(modl2)
Start:  AIC=451.85
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize +
    bank:provincesa + lotsize:provincesa

                     Df Deviance    AIC
- bank:provincesa     6   415.15 445.15
- lotsize:provincesa  3   410.63 446.63
- bank:lotsize        2   410.91 448.91
<none>                    409.85 451.85
- bank:time           2   421.92 459.92

Step:  AIC=445.15
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize +
    lotsize:provincesa

                     Df Deviance    AIC
- lotsize:provincesa  3   415.83 439.83
- bank:lotsize        2   416.09 442.09
<none>                    415.15 445.15
- bank:time           2   431.51 457.51
```

```
Step:  AIC=439.83
yield ~ bank + time + lotsize + provincesa + bank:time + bank:lotsize

               Df Deviance    AIC
- bank:lotsize  2   416.91 436.91
<none>              415.83 439.83
- bank:time     2   432.01 452.01
- provincesa    3   457.71 475.71

Step:  AIC=436.91
yield ~ bank + time + lotsize + provincesa + bank:time

             Df Deviance    AIC
<none>            416.91 436.91
- lotsize     1   423.76 441.76
- bank:time   2   445.68 461.68
- provincesa  3   458.57 472.57
Warning messages:
1: In glm.fit(x[, jj, drop = FALSE], y, wt, offset = object$offset,  :
  fitted probabilities numerically 0 or 1 occurred
2: In glm.fit(x[, jj, drop = FALSE], y, wt, offset = object$offset,  :
  fitted probabilities numerically 0 or 1 occurred
3: In glm.fit(x[, jj, drop = FALSE], y, wt, offset = object$offset,  :
  fitted probabilities numerically 0 or 1 occurred
>
> writeLines("\n\n\nANOVA Table for Simplified Model\n\n\n")



ANOVA Table for Simplified Model



> anova(modl3,test="Chisq")
Analysis of Deviance Table

Model: binomial, link: logit

Response: yield

Terms added sequentially (first to last)


           Df Deviance Resid. Df Resid. Dev P(>|Chi|)
NULL                         690     953.86
bank        2  253.745       688     700.12 < 2.2e-16 ***
time        1  211.606       687     488.51 < 2.2e-16 ***
lotsize     1    3.285       686     485.22    0.0699 .
provincesa  3   39.545       683     445.68 1.330e-08 ***
bank:time   2   28.771       681     416.91 5.657e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> writeLines("\n\n\nTest significance of lotsize term
explicitly\n\n\n")



Test significance of lotsize term explicitly



> modl3a<-update(modl3,~.-lotsize)
> anova(modl3,modl3a,test="Chisq")
Analysis of Deviance Table

Model 1: yield ~ bank + time + lotsize + provincesa + bank:time
Model 2: yield ~ bank + time + provincesa + bank:time
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
```

```
1       681     416.91
2       682     423.76 -1  -6.8523  0.008853 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> writeLines("\n\n\nLotsize is significant at 1% level so retain in
model\n\n\n")



Lotsize is significant at 1% level so retain in model



>
>
#***********************************************************************
*********
> #
> #    Confusion Matrix for Simplified Model after excluding third
order term
> #
>
#***********************************************************************
*********
>
> writeLines("\n\n\nConfusion Matrix for Simplified Model p=0.5\n\n\n")



Confusion Matrix for Simplified Model p=0.5



>
> pred<-(rep("Low",length(fitted(modl3))))
> pred<-factor(pred,levels=c("Low","High"))
> pred[fitted(modl3)<0.5]<-"Low"
> pred[fitted(modl3)>=0.5]<-"High"
> table(pred,yield=opdata$yield)
      yield
pred   Low High
  Low  287   12
  High  85  307
>
>
>
#***********************************************************************
*************************
> #
> #     Co-efficients for fitted Simplified Model
> #
>
#***********************************************************************
*************************
> writeLines("\n\n\nCo-efficients for Simplified Model\n\n\n")



Co-efficients for Simplified Model



> coef(modl3)
        (Intercept)          bankHSBC            bankLTSB
         15.7818902       -12.0676382         -14.0907971
           timeLate        lotsizeLarge   provincesaMidlands
        -20.9331562         0.7759436          -0.8785792
     provincesaSouth      provincesaCelt bankHSBC   :timeLate
         -1.5737788         0.9819891          13.9879102
```

```
bankLTSB   :timeLate
          17.8488444
> anova(modl3,test="Chisq")
Analysis of Deviance Table

Model: binomial, link: logit

Response: yield

Terms added sequentially (first to last)


           Df Deviance Resid. Df Resid. Dev P(>|Chi|)
NULL                         690     953.86
bank        2  253.745       688     700.12 < 2.2e-16 ***
time        1  211.606       687     488.51 < 2.2e-16 ***
lotsize     1    3.285       686     485.22    0.0699 .
provincesa  3   39.545       683     445.68 1.330e-08 ***
bank:time   2   28.771       681     416.91 5.657e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> summary(modl3)

Call:
glm(formula = yield ~ bank + time + lotsize + provincesa + bank:time,
    family = binomial, data = opdata)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-3.00045  -0.18782  -0.07218   0.47156   3.45013

Coefficients:
                     Estimate Std. Error z value Pr(>|z|)
(Intercept)           15.7819   700.1381   0.023 0.982016
bankHSBC             -12.0676   700.1381  -0.017 0.986248
bankLTSB             -14.0908   700.1380  -0.020 0.983943
timeLate             -20.9332   700.1387  -0.030 0.976148
lotsizeLarge           0.7759     0.3067   2.530 0.011396 *
provincesaMidlands    -0.8786     0.4446  -1.976 0.048145 *
provincesaSouth       -1.5738     0.4144  -3.797 0.000146 ***
provincesaCelt         0.9820     0.6159   1.594 0.110835
bankHSBC   :timeLate  13.9879   700.1392   0.020 0.984060
bankLTSB   :timeLate  17.8488   700.1388   0.025 0.979661
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 953.86  on 690  degrees of freedom
Residual deviance: 416.91  on 681  degrees of freedom
AIC: 436.91

Number of Fisher Scoring iterations: 14

>
>
#**********************************************************************
************************
> #
> #     Cross Validation of Simplified Model with data divided randomly
into 10 approx.equal groups
> #
>
#**********************************************************************
************************
>
>
>
>
```

```
>
#*********************************************************************
************************
> #     Randomly divide data into 10 equal groups
>
#*********************************************************************
************************
> randnos<-with(opdata,runif(length(yield)))
> orderednos<-order(randnos)
> full.length<-with(opdata,full.length<-(length(yield)))
> sub.length<-floor(full.length/10)
> sections<-seq(0,full.length,sub.length)
> sections[11]<-full.length
>
>
>
#*********************************************************************
************************
> #     Cross Validate
>
#*********************************************************************
************************
>
> model.formula<-formula(modl3)
>
> cross.val.data<-cbind(opdata,randnos,orderednos)
>
>
> for(i in 1:10){
+ cross.val.sample<-with(cross.val.data,{
+ tempdata0<-cross.val.data[{orderednos<=sections[i] |
orderednos>sections[i+1]},]
+ return(tempdata0)
+ })
+ cross.val.test<-with(cross.val.data,{
+ tempdata1<-cross.val.data[{orderednos>sections[i] &
orderednos<=sections[i+1]},]
+ return(tempdata1)
+ })
+
+ fitted.model<-glm(model.formula,family=binomial,cross.val.sample)
+
+ fitted.values<-
predict(fitted.model,newdata=cross.val.test,type="response")
+ outfile<-cbind(cross.val.test,fitted=fitted.values)
+
+ if(i==1)results<-outfile
+ else results<-rbind(results,outfile)
+
+ }
>
>
>
>
#*********************************************************************
************************
> #
> #     Confusion Matrix for Simplified Model after Cross Validation
> #
>
#*********************************************************************
************************
> writeLines("\n\n\nConfusion Matrix after Cross Validation -
Simplified Model p=0.5\n\n\n")



Confusion Matrix after Cross Validation - Simplified Model p=0.5
```

```
>
> pred<-(rep("Low",length(results$fitted)))
> pred<-factor(pred,levels=c("Low","High"))
> pred[results$fitted<0.5]<-"Low"
> pred[results$fitted>=0.5]<-"High"
> table(pred,yield=results$yield)
      yield
pred   Low High
  Low  287   12
  High  85  307
>
>
>
>
#*********************************************************************
************************
> #
> #     ROC chart for Simplified Model after Cross Validation
> #
>
#*********************************************************************
************************
> writeLines("\n\n\nROC Chart after Cross Validation - Simplified
Model\n\n\n")



ROC Chart after Cross Validation - Simplified Model



>
> simple.pred<-
prediction(results$fitted,results$yield,label.ordering=c("Low","High"))
> simple.perf<-performance(simple.pred,"tpr","fpr")
> plot(simple.perf,main="ROC Chart after Cross Validation - Simplified
Model Provinces A",cex.main=0.9,colorize=TRUE)
> text(locator(1),"Key to Cutoff Probability",srt=90)
>
> perf<-performance(simple.pred,"prec","fpr")
> plot(perf,main="Precision versus False Positive Error Rate Provinces
A",cex.main=0.9,
+             sub="Simplified Model",cex.sub=0.9,
+             colorize=T)
> text(locator(1),"Key to Cutoff Probability",srt=90)
>
>
>
#*********************************************************************
************************
> #     Output ROC chart data for Simplified Model after Cross
Validation
>
#*********************************************************************
************************
> rocout<-data.frame(fitted=results$fitted,yield=results$yield)
> write.table(rocout,file="provasimpleroc.out")
>
>
>
#*********************************************************************
*******
> #
> #                          bar chart
> #
```

```
>
#*********************************************************************
*******
>
> limits<-seq(0.0,1.0,0.1)
> temppred<-rep(1,length(results$fitted))
>
> predgrp<-with(results,{
+ for(i in 1:10)temppred[fitted>limits[i] & fitted<=limits[i+1]]<-i
+ return(temppred)
+ }
+ )
>
> dummy<-rep(1,length(predgrp))
> add.results<-cbind(results,pred.grp=predgrp,counter=dummy)
>
>
> #"barchart.data" is the array containing the number of low yield and
high yield observations
> # whose prdicted probabilites of being in thehigh yield group are in
the ranges
> #specified by "intervals"
>
> barchart.data<-
with(add.results,tapply(counter,list(yield,pred.grp),sum))
>
> intervals<-c("0.00-0.09","0.10-0.19","0.20-0.29","0.30-0.39","0.40-
0.49",
+  "0.50-0.59","0.60-0.69","0.70-0.79","0.80-0.89","0.90-1.00")
>
>
barplot(barchart.data,beside=TRUE,names.arg=intervals,space=c(1,5),cex.
names=0.5,
+   main="Counts against Probability(high yield) Provinces
A",cex.main=0.9,
+         sub="Simplified Model",cex.sub=0.9,
+   xlab="Probability case belongs to high yield group",
+   ylab="No of cases",
+   col=c("grey25","grey75"))
>
> legend(locator(1),legend=c("low yield","high
yield"),fill=c("grey25","grey75"))
>
>
>
>
#*********************************************************************
*******
> #
> #         ROC Chart comparing Basic and Simplified Models
> #
>
#*********************************************************************
*******
> writeLines("\n\n\nROC Chart comparing Basic and Simplified
Models\n\n\n")



ROC Chart comparing Basic and Simplified Models



>
>
> basic.perf<-performance(basic.pred,"tpr","fpr")
> plot(basic.perf,col="blue",main="ROC Chart comparing Basic and
Simplified Models Provinces A",cex.main=0.9)
> plot(simple.perf,col="red",add=TRUE)
```

```
> legend(locator(1),legend=c("basic model","simplified
model"),fill=c("blue","red"))
>
```

Appendix VIII

Provinces B logistic regression output

```
#
> # Malvern Tipping logistic cross validation Analysis in R
> #
>
>
>
#*********************************************************************
************************
> #
> #            remove "#"s to output text to file
> #
> #    NB To print out figures must still step through script using
CTRL-R so have time to
> #    access "file" header to output graphs to file
> #
>
#*********************************************************************
************************
> #sink(file="logisticcrossvalidationProvB.txt")
>
>
>
#*********************************************************************
************************
> #
> # Load package 'foreign' to access read.spss function
> #
#*********************************************************************
************************
> library(foreign)
>
>
>
>
#*********************************************************************
************************
> #
> #         read data from an SPSS file and convert to a data.frame
> #
>
#*********************************************************************
************************
> filename<-"Three Main Banks C.sav"
> RawData<-read.spss(filename,to.data.frame=TRUE)
>
> writeLines("\n\n\nNames of variables in Input File\n\n\n")



Names of variables in Input File



> names(RawData)
 [1] "CaseNo"     "Date"       "Bank"       "Region"     "Location"
 [6] "Price"      "Rent"       "Yield"      "Lotsize"    "Tenure"
[11] "Quarter"    "BankNo"     "RegCode"    "SizeCode"   "TenuCode"
[16] "Year"       "ProvincesB" "YieldGroup" "LotGroup"   "Time"
>
> writeLines("\n\n\nSummary of Input Data\n\n\n")
```

Summary of Input Data

```
> summary(RawData,maxsum=25)
     CaseNo                    Date             Bank                 Region   
 Min.   :  1.0    15 Jul 1997: 97   Barclay:192   East Anglia: 32  
 1st Qu.:254.5    27 Jun 2001: 92   HSBC   :160   East Mid   : 53  
 Median :520.0    21 Nov 2001: 54   LTSB   :339   London-M25 : 49  
 Mean   :504.1    29 Mar 2006: 34                 North-East : 61  
 3rd Qu.:762.5    05 Jul 2006: 25                 North-West : 53  
 Max.   :968.0    24 May 2006: 24                 Scotland   : 15  
                  16 May 2006: 23                 South-East :193  
                  06 Jul 2006: 21                 South-West :103  
                  06 Jul 2005: 20                 Wales      : 65  
                  03 Jul 2006: 19                 West Mid   : 67  
                  28 Mar 2006: 19                                  
                  17 Oct 2001: 17                                  
                  22 Feb 2006: 16                                  
                  19 Mar 2002: 13                                  
                  22 Mar 2006: 13                                  
                  08 Oct 2003: 12                                  
                  10 Feb 2004: 12                                  
                  19 May 2004: 11                                  
                  22 May 2001: 11                                  
                  25 May 2006: 11                                  
                  23 May 2002: 10                                  
                  25 Mar 2004: 10                                  
                  04 Feb 2002:  9                                  
                  03 Dec 2003:  8                                  
                  (Other)    :110                                  
                        Location        Price               Rent       
 Chatham, Kent              :  4   Min.   :  69000   Min.   :  4400  
 Melton Mowbray, Leics      :  4   1st Qu.: 250000   1st Qu.: 16500  
 Tadley, Hampshire          :  4   Median : 463000   Median : 27000  
 Brixham, Devon             :  3   Mean   : 650092   Mean   : 37262  
 Coulsdon, Surrey           :  3   3rd Qu.: 810000   3rd Qu.: 45000  
 Crediton, Devon            :  3   Max.   :5310000   Max.   :275000  
 Cullumpton, Devon          :  3                                     
 East Ham, London           :  3                                     
 Maldon, Essex              :  3                                     
 Okehampton, Devon          :  3                                     
 Wareham, Dorset            :  3                                     
 Amesbury, Wiltshire        :  2                                     
 Ammanford, Dyfed           :  2                                     
 Aylsham, Norfolk           :  2                                     
 Bala, Gwynedd              :  2                                     
 Bedford, Beds              :  2                                     
 Bedminster, Avon           :  2                                     
 Bicester, Oxfordshire      :  2                                     
 Bodmin, Cornwall           :  2                                     
 Boscombe, Bournemouth      :  2                                     
 Bridgwater, Somerset       :  2                                     
 Bromyard, Herefordshire :  2                                     
 Camberley, Surrey          :  2                                     
 Cleckheaton, West Yorks :  2                                     
 (Other)                    :629                                     
     Yield           Lotsize          Tenure          Quarter    
 Min.   :3.170    L       :316   F       :679   Min.   :1997  
 1st Qu.:5.040    M       :200   Feudal  :  2   1st Qu.:2001  
 Median :6.000    S       :175   Feuhold : 10   Median :2002  
 Mean   :6.223                                  Mean   :2003  
 3rd Qu.:7.460                                  3rd Qu.:2006  
 Max.   :9.900                                  Max.   :2006  
```

```
                    BankNo             RegCode      SizeCode
TenuCode
 hsbc                  :160   London M25   : 48   small :173   F
:679
 lloydstsb             :339   South east   :194   medium:197   feuhold
: 12
 natwest               :  0   South west   :103   large :321   L
:  0
 rbs                   :  0   East anglia  : 32
Heritable:  0
 BW                    :  0   West midlands: 67
 BB                    :  0   East Midlands: 53
 barclays              :192   North east   : 61
 yorkshire             :  0   North west   : 53
 woolwich              :  0   Wales        : 65
 cheltenham and glos   :  0   Scotland     : 15
 abbey national        :  0
 clydesdale            :  0
 halifax/bos           :  0
 alliance and leicester:  0
```

```
      Year               ProvincesB  YieldGroup  LotGroup      Time
 Min.   :1997   London & SE  :242   Low :372   Small:370   Early:391
 1st Qu.:2001   Wales & SW   :168   High:319   Large:321   Late :300
 Median :2002   Midlands     :152
 Mean   :2003   North Britain:129
 3rd Qu.:2006
 Max.   :2006
```

```
>
>
#######################################################################
########################
> #
> #     Extract required data into Operating Data File
> #
>
#######################################################################
########################
>
>
> opdata<-with(RawData,{
+ bank<-factor(Bank)
+ time<-factor(Time)
+ provincesb<-factor(ProvincesB)
+ yield<-factor(YieldGroup)
+ lotsize<-factor(LotGroup)
+ oplist<-list(yield=as.factor(yield),bank=as.factor(bank),
+  time=as.factor(time),lotsize=as.factor(lotsize),
+  provincesb=as.factor(provincesb))
+ opdata<-as.data.frame(oplist)
+ return(opdata)})
>
> writeLines("\n\n\nSummary of Operating Data used in Analysis\n\n\n")



Summary of Operating Data used in Analysis



> summary(opdata)
  yield         bank        time       lotsize           provincesb
 Low :372   Barclay:192   Early:391   Small:370   London & SE  :242
 High:319   HSBC   :160   Late :300   Large:321   Wales & SW   :168
            LTSB   :339                           Midlands     :152
                                                  North Britain:129
>
>
>
>
#######################################################################
########################
> #
> #     Basic Model using step for backwards elimination
> #
>
#######################################################################
########################
>
> modl0<-
glm(yield~(bank+time+lotsize+provincesb)^4,family=binomial,opdata)
>
> writeLines("\n\n\nBasic Model after backwards elimination from
Saturated Model\nusing Akaike's Information Criterion\n\n\n")



Basic Model after backwards elimination from Saturated Model
using Akaike's Information Criterion



> modl1<-step(modl0)
Start:  AIC=483.91
```

```
yield ~ (bank + time + lotsize + provincesb)^4

                              Df Deviance    AIC
- bank:time:lotsize:provincesb  2   401.91 479.91
<none>                              401.91 483.91

Step:  AIC=479.91
yield ~ bank + time + lotsize + provincesb + bank:time + bank:lotsize +
    bank:provincesb + time:lotsize + time:provincesb +
lotsize:provincesb +
    bank:time:lotsize + bank:time:provincesb + bank:lotsize:provincesb
+
    time:lotsize:provincesb

                         Df Deviance    AIC
- bank:time:provincesb     5   401.91 469.91
- bank:time:lotsize        1   401.91 477.91
<none>                         401.91 479.91
- time:lotsize:provincesb  3   407.95 479.95
- bank:lotsize:provincesb  5   412.24 480.24

Step:  AIC=469.91
yield ~ bank + time + lotsize + provincesb + bank:time + bank:lotsize +
    bank:provincesb + time:lotsize + time:provincesb +
lotsize:provincesb +
    bank:time:lotsize + bank:lotsize:provincesb +
time:lotsize:provincesb

                         Df Deviance    AIC
- bank:time:lotsize        1   401.91 467.91
<none>                         401.91 469.91
- time:lotsize:provincesb  3   407.95 469.95
- bank:lotsize:provincesb  5   418.00 476.00

Step:  AIC=467.91
yield ~ bank + time + lotsize + provincesb + bank:time + bank:lotsize +
    bank:provincesb + time:lotsize + time:provincesb +
lotsize:provincesb +
    bank:lotsize:provincesb + time:lotsize:provincesb

                         Df Deviance    AIC
<none>                         401.91 467.91
- time:lotsize:provincesb  3   407.95 467.95
- bank:time                2   410.79 472.79
- bank:lotsize:provincesb  5   419.20 475.20
There were 14 warnings (use warnings() to see them)
>
> anova(modl1,test="Chisq")
Analysis of Deviance Table

Model: binomial, link: logit

Response: yield

Terms added sequentially (first to last)


                       Df Deviance Resid. Df Resid. Dev P(>|Chi|)
NULL                                     690     953.86
bank                    2  253.745       688     700.12 < 2.2e-16 ***
time                    1  211.606       687     488.51 < 2.2e-16 ***
lotsize                 1    3.285       686     485.22  0.069906 .
provincesb              3   21.403       683     463.82 8.682e-05 ***
bank:time               2   26.606       681     437.21 1.669e-06 ***
bank:lotsize            2    0.861       679     436.35  0.650244
bank:provincesb         6    6.846       673     429.51  0.335345
time:lotsize            1    1.012       672     428.50  0.314501
time:provincesb         3    2.774       669     425.72  0.427753
lotsize:provincesb      3    2.683       666     423.04  0.443162
```

```
bank:lotsize:provincesb  5   15.090       661     407.95  0.009985 **
time:lotsize:provincesb  3    6.039       658     401.91  0.109746
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Warning messages:
1: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
2: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
3: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
4: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
5: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
>
> writeLines("\n\n\nSignificance Test for bank:lotsize:provincesB
term\n\n\n")



Significance Test for bank:lotsize:provincesB term



> modl2a<-
glm(yield~(bank+time+lotsize+provincesb)^2+time:lotsize:provincesb,fami
ly=binomial,opdata)
Warning message:
In glm.fit(x = X, y = Y, weights = weights, start = start, etastart =
etastart,  :
  fitted probabilities numerically 0 or 1 occurred
> anova(modl1,modl2a,test="Chi")
Analysis of Deviance Table

Model 1: yield ~ bank + time + lotsize + provincesb + bank:time +
bank:lotsize +
    bank:provincesb + time:lotsize + time:provincesb +
lotsize:provincesb +
    bank:lotsize:provincesb + time:lotsize:provincesb
Model 2: yield ~ (bank + time + lotsize + provincesb)^2 +
time:lotsize:provincesb
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       658     401.91
2       663     419.20 -5  -17.292  0.003978 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> writeLines("\n\n\nSignificance Test for time:lotsize:provincesB
term\n\n\n")



Significance Test for time:lotsize:provincesB term



> modl2b<-
glm(yield~(bank+time+lotsize+provincesb)^2+bank:lotsize:provincesb,fami
ly=binomial,opdata)
Warning message:
In glm.fit(x = X, y = Y, weights = weights, start = start, etastart =
etastart,  :
  fitted probabilities numerically 0 or 1 occurred
```

```
> anova(modl1,modl2b,test="Chi")
Analysis of Deviance Table

Model 1: yield ~ bank + time + lotsize + provincesb + bank:time +
bank:lotsize +
    bank:provincesb + time:lotsize + time:provincesb +
lotsize:provincesb +
    bank:lotsize:provincesb + time:lotsize:provincesb
Model 2: yield ~ (bank + time + lotsize + provincesb)^2 +
bank:lotsize:provincesb
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       658     401.91
2       661     407.95 -3  -6.0386    0.1097
>
> writeLines("\n\n\nChoose model 2b as the Basic Model excluding
time:lotsize:ProvincesB term\n\n\n")



Choose model 2b as the Basic Model excluding time:lotsize:ProvincesB
term



>
>
>
#**********************************************************************
*************************
> #
> #     Confusion Matrix for Basic Model p=0.5
> #
>
#**********************************************************************
*************************
> writeLines("\n\n\nConfusion Matrix for Basic Model p=0.5\n\n\n")



Confusion Matrix for Basic Model p=0.5



>
> pred<-(rep("Low",length(fitted(modl2b))))
> pred<-factor(pred,levels=c("Low","High"))
> pred[fitted(modl2b)<0.5]<-"Low"
> pred[fitted(modl2b)>=0.5]<-"High"
> table(pred,yield=opdata$yield)
      yield
pred   Low High
  Low  284   10
  High  88  309
>
>
>
>
#**********************************************************************
*************************
> #
> #     Co-efficients for fitted Basic Model
> #
>
#**********************************************************************
*************************
> writeLines("\n\n\nCo-efficients for Basic Model\n\n\n")
```

Co-efficients for Basic Model

```
> coef(modl2b)
                                    (Intercept)
                                     19.9800341
                                    bankHSBC
                                    -17.8205499
                                    bankLTSB
                                    -19.4839811
                                       timeLate
                                    -39.5461027
                                   lotsizeLarge
                                     14.7593404
                          provincesbWales & SW
                                     16.2115396
                            provincesbMidlands
                                     -0.5843464
                       provincesbNorth Britain
                                    -16.5079428
                          bankHSBC   :timeLate
                                      2.7011434
                          bankLTSB   :timeLate
                                     35.6244367
                      bankHSBC   :lotsizeLarge
                                     17.9526199
                      bankLTSB   :lotsizeLarge
                                    -14.6591409
              bankHSBC   :provincesbWales & SW
                                    -15.6194885
              bankLTSB   :provincesbWales & SW
                                    -16.2375890
                bankHSBC   :provincesbMidlands
                                     17.3929154
                bankLTSB   :provincesbMidlands
                                      0.8191966
           bankHSBC   :provincesbNorth Britain
                                     33.2545129
           bankLTSB   :provincesbNorth Britain
                                     18.2657195
                         timeLate:lotsizeLarge
                                      0.3640769
                 timeLate:provincesbWales & SW
                                    -16.2115396
                   timeLate:provincesbMidlands
                                      0.5843464
              timeLate:provincesbNorth Britain
                                      1.0050087
             lotsizeLarge:provincesbWales & SW
                                    -15.2024691
               lotsizeLarge:provincesbMidlands
                                    -15.1234173
          lotsizeLarge:provincesbNorth Britain
                                      0.4625724
   bankHSBC   :lotsizeLarge:provincesbWales & SW
                                     12.3330737
   bankLTSB   :lotsizeLarge:provincesbWales & SW
                                     16.2417039
     bankHSBC   :lotsizeLarge:provincesbMidlands
                                    -20.2261288
     bankLTSB   :lotsizeLarge:provincesbMidlands
                                     16.1401988
bankHSBC   :lotsizeLarge:provincesbNorth Britain
                                    -68.9542233
bankLTSB   :lotsizeLarge:provincesbNorth Britain
                                             NA
> anova(modl2b,test="Chisq")
Analysis of Deviance Table
```

```
Model: binomial, link: logit

Response: yield

Terms added sequentially (first to last)


                        Df Deviance Resid. Df Resid. Dev P(>|Chi|)
NULL                                      690     953.86
bank                     2  253.745       688     700.12 < 2.2e-16 ***
time                     1  211.606       687     488.51 < 2.2e-16 ***
lotsize                  1    3.285       686     485.22  0.069906 .
provincesb               3   21.403       683     463.82 8.682e-05 ***
bank:time                2   26.606       681     437.21 1.669e-06 ***
bank:lotsize             2    0.861       679     436.35  0.650244
bank:provincesb          6    6.846       673     429.51  0.335345
time:lotsize             1    1.012       672     428.50  0.314501
time:provincesb          3    2.774       669     425.72  0.427753
lotsize:provincesb       3    2.683       666     423.04  0.443162
bank:lotsize:provincesb  5   15.090       661     407.95  0.009985 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Warning messages:
1: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
2: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
3: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
4: In method(x = x[, varseq <= i, drop = FALSE], y = y, weights =
object$prior.weights,  :
  fitted probabilities numerically 0 or 1 occurred
> summary(modl2b)

Call:
glm(formula = yield ~ (bank + time + lotsize + provincesb)^2 +
    bank:lotsize:provincesb, family = binomial, data = opdata)

Deviance Residuals:
       Min          1Q      Median          3Q         Max
-2.372e+00  -1.529e-01  -7.598e-05   4.673e-01   2.985e+00

Coefficients: (1 not defined because of singularities)
                                                  Estimate Std. Error
(Intercept)                                        19.9800  6183.6000
bankHSBC                                          -17.8205  6183.6000
bankLTSB                                          -19.4840  6183.6000
timeLate                                          -39.5461  5265.2998
lotsizeLarge                                       14.7593  3242.4572
provincesbWales & SW                               16.2115  8558.1312
provincesbMidlands                                 -0.5843  7004.5076
provincesbNorth Britain                           -16.5079  1883.5400
bankHSBC   :timeLate                                2.7011  5713.7633
bankLTSB   :timeLate                               35.6244  5265.2996
bankHSBC   :lotsizeLarge                           17.9526  3929.0255
bankLTSB   :lotsizeLarge                          -14.6591  3242.4573
bankHSBC   :provincesbWales & SW                  -15.6195  8558.1313
bankLTSB   :provincesbWales & SW                  -16.2376  8558.1312
bankHSBC   :provincesbMidlands                     17.3929  7360.7419
bankLTSB   :provincesbMidlands                      0.8192  7004.5076
bankHSBC   :provincesbNorth Britain                33.2545  2407.5199
bankLTSB   :provincesbNorth Britain                18.2657  1883.5399
timeLate:lotsizeLarge                               0.3641     1.0509
timeLate:provincesbWales & SW                     -16.2115  2214.4267
timeLate:provincesbMidlands                         0.5843     1.3420
```

```
timeLate:provincesbNorth Britain                     1.0050     1.3480
lotsizeLarge:provincesbWales & SW                  -15.2025  8700.5978
lotsizeLarge:provincesbMidlands                    -15.1234  7213.0770
lotsizeLarge:provincesbNorth Britain                 0.4626     1.0863
bankHSBC   :lotsizeLarge:provincesbWales & SW       12.3331 11394.5058
bankLTSB   :lotsizeLarge:provincesbWales & SW       16.2417  8700.5979
bankHSBC   :lotsizeLarge:provincesbMidlands        -20.2261  8289.1951
bankLTSB   :lotsizeLarge:provincesbMidlands         16.1402  7213.0771
bankHSBC   :lotsizeLarge:provincesbNorth Britain   -68.9542  4323.3980
bankLTSB   :lotsizeLarge:provincesbNorth Britain         NA         NA
                                                    z value Pr(>|z|)
(Intercept)                                           0.003    0.997
bankHSBC                                             -0.003    0.998
bankLTSB                                             -0.003    0.997
timeLate                                             -0.008    0.994
lotsizeLarge                                          0.005    0.996
provincesbWales & SW                                  0.002    0.998
provincesbMidlands                                -8.34e-05    1.000
provincesbNorth Britain                              -0.009    0.993
bankHSBC   :timeLate                               4.73e-04    1.000
bankLTSB   :timeLate                                  0.007    0.995
bankHSBC   :lotsizeLarge                              0.005    0.996
bankLTSB   :lotsizeLarge                             -0.005    0.996
bankHSBC   :provincesbWales & SW                     -0.002    0.999
bankLTSB   :provincesbWales & SW                     -0.002    0.998
bankHSBC   :provincesbMidlands                        0.002    0.998
bankLTSB   :provincesbMidlands                     1.17e-04    1.000
bankHSBC   :provincesbNorth Britain                   0.014    0.989
bankLTSB   :provincesbNorth Britain                   0.010    0.992
timeLate:lotsizeLarge                                 0.346    0.729
timeLate:provincesbWales & SW                        -0.007    0.994
timeLate:provincesbMidlands                           0.435    0.663
timeLate:provincesbNorth Britain                      0.746    0.456
lotsizeLarge:provincesbWales & SW                    -0.002    0.999
lotsizeLarge:provincesbMidlands                      -0.002    0.998
lotsizeLarge:provincesbNorth Britain                  0.426    0.670
bankHSBC   :lotsizeLarge:provincesbWales & SW         0.001    0.999
bankLTSB   :lotsizeLarge:provincesbWales & SW         0.002    0.999
bankHSBC   :lotsizeLarge:provincesbMidlands          -0.002    0.998
bankLTSB   :lotsizeLarge:provincesbMidlands           0.002    0.998
bankHSBC   :lotsizeLarge:provincesbNorth Britain     -0.016    0.987
bankLTSB   :lotsizeLarge:provincesbNorth Britain         NA       NA

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 953.86  on 690  degrees of freedom
Residual deviance: 407.95  on 661  degrees of freedom
AIC: 467.95

Number of Fisher Scoring iterations: 18

>
>
>
>
>
#***********************************************************************
************************
> #
> #     Cross Validation of Basic Model with data divided randomly into
10 approx.equal groups
> #
>
#***********************************************************************
************************
>
>
>
>
```

```
>
#***********************************************************************
************************
> #     Randomly divide data into 10 equal groups
>
#***********************************************************************
************************
>
> randnos<-with(opdata,runif(length(yield)))
> orderednos<-order(randnos)
> full.length<-with(opdata,full.length<-(length(yield)))
> sub.length<-floor(full.length/10)
> sections<-seq(0,full.length,sub.length)
> sections[11]<-full.length
>
>
>
#***********************************************************************
************************
> #     Cross Validate
>
#***********************************************************************
************************
> model.formula<-formula(modl2b)
> cross.val.data<-cbind(opdata,randnos,orderednos)
>
>
> for(i in 1:10){
+ cross.val.sample<-with(cross.val.data,{
+ tempdata0<-cross.val.data[{orderednos<=sections[i] |
orderednos>sections[i+1]},]
+ return(tempdata0)
+ })
+
+ cross.val.test<-with(cross.val.data,{
+ tempdata1<-cross.val.data[{orderednos>sections[i] &
orderednos<=sections[i+1]},]
+ return(tempdata1)
+ })
+
+ fitted.model<-glm(model.formula,family=binomial,cross.val.sample)
+
+ fitted.values<-
predict(fitted.model,newdata=cross.val.test,type="response")
+ outfile<-cbind(cross.val.test,fitted=fitted.values)
+
+ if(i==1)results<-outfile
+ else results<-rbind(results,outfile)
+
+
+ }
There were 20 warnings (use warnings() to see them)
>
>
>
#***********************************************************************
************************
> #
> #     Confusion Matrix for Basic Model after Cross Validation
> #
>
#***********************************************************************
************************
> writeLines("\n\n\nConfusion Matrix after Cross Validation - Basic
Model p=0.5\n\n\n")



Confusion Matrix after Cross Validation - Basic Model p=0.5
```

```
>
> pred<-(rep("Low",length(results$fitted)))
> pred<-factor(pred,levels=c("Low","High"))
> pred[results$fitted<0.5]<-"Low"
> pred[results$fitted>=0.5]<-"High"
>
> table(pred,yield=results$yield)
      yield
pred   Low High
  Low  283   12
  High  89  307
>
>
>
#***********************************************************************
************************
> #
> #     ROC chart for Basic Model after Cross Validation
> #
>
#***********************************************************************
************************
> writeLines("\n\n\nROC Chart after Cross Validation - Basic
Model\n\n\n")



ROC Chart after Cross Validation - Basic Model



>
> library(ROCR)
> basic.pred<-
prediction(results$fitted,results$yield,label.ordering=c("Low","High"))
> perf<-performance(basic.pred,"tpr","fpr")
> plot(perf,main="ROC Chart after Cross Validation - Basic Model
Provinces B",cex.main=0.9,colorize=TRUE)
> text(locator(1),"Key to Cutoff Probability",srt=90)
>
>
>
#***********************************************************************
************************
> #     Output ROC chart data for Basic Model after Cross Validation
>
#***********************************************************************
************************
> rocout<-data.frame(fitted=results$fitted,yield=results$yield)
> write.table(rocout,file="provbbasicroc.out")
>
> #NB ROC data is also sufficient to reproduce the bar chart.
>
>
#***********************************************************************
*******
> #
> #                           bar chart
> #
>
#***********************************************************************
*******
>
> limits<-seq(0.0,1.0,0.1)
> temppred<-rep(1,length(results$fitted))
>
```

```
> predgrp<-with(results,{
+ for(i in 1:10)temppred[fitted>limits[i] & fitted<=limits[i+1]]<-i
+ return(temppred)
+ }
+ )
>
> dummy<-rep(1,length(predgrp))
> add.results<-cbind(results,pred.grp=predgrp,counter=dummy)
>
> #"data" is the array containing the number of low yield and high
yield observations
> # whose prdicted probabilites of being in thehigh yield group are in
the ranges
> #specified by "intervals"
>
> barchart.data<-
with(add.results,tapply(counter,list(yield,pred.grp),sum))
>
> intervals<-c("0.00-0.09","0.10-0.19","0.20-0.29","0.30-0.39","0.40-
0.49",
+  "0.50-0.59","0.60-0.69","0.70-0.79","0.80-0.89","0.90-1.00")
>
>
barplot(barchart.data,beside=TRUE,names.arg=intervals,space=c(1,5),cex.
names=0.5,
+   xlab="Probability case belongs to high yield group",
+   ylab="No of cases",
+   col=c("grey25","grey75"),
+   main="Counts against Probability(high yield) Provinces
B",cex.main=0.9,
+         sub="Basic Model",cex.sub=0.9)
>
> legend(locator(1),legend=c("low yield","high
yield"),fill=c("grey25","grey75"))
>
>
>
>
>
#**********************************************************************
*********
> #
> #    Simplified model excluding higher order term
> #
>
#**********************************************************************
*********
>
> writeLines("\n\n\nExamine STATISTICAL significance of third order
term bank:lotsize:provinces\n\n\n")



Examine STATISTICAL significance of third order term
bank:lotsize:provinces



> modl3<-update(modl2b,~.-bank:lotsize:provincesb)
Warning message:
In glm.fit(x = X, y = Y, weights = weights, start = start, etastart =
etastart,  :
  fitted probabilities numerically 0 or 1 occurred
> anova(modl2b,modl3,test="Chisq")
Analysis of Deviance Table

Model 1: yield ~ (bank + time + lotsize + provincesb)^2 +
bank:lotsize:provincesb
```

```
Model 2: yield ~ bank + time + lotsize + provincesb + bank:time +
bank:lotsize +
    bank:provincesb + time:lotsize + time:provincesb +
lotsize:provincesb
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       661     407.95
2       666     423.04 -5   -15.09  0.009985 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
>
> writeLines("\n\n\nAlthough term is STATISTICALLY significant remove
it and see if it has PRACTICAL significance\n\n\n")



Although term is STATISTICALLY significant remove it and see if it has
PRACTICAL significance



> modl4<-step(modl3)
Start:  AIC=473.04
yield ~ bank + time + lotsize + provincesb + bank:time + bank:lotsize +
    bank:provincesb + time:lotsize + time:provincesb +
lotsize:provincesb

                     Df Deviance    AIC
- bank:provincesb     6   429.55 467.55
- lotsize:provincesb  3   425.72 469.72
- bank:lotsize        2   424.19 470.19
- time:provincesb     3   426.90 470.90
- time:lotsize        1   424.79 472.79
<none>                    423.04 473.04
- bank:time           2   433.76 479.76

Step:  AIC=467.55
yield ~ bank + time + lotsize + provincesb + bank:time + bank:lotsize +
    time:lotsize + time:provincesb + lotsize:provincesb

                     Df Deviance    AIC
- lotsize:provincesb  3   432.10 464.10
- bank:lotsize        2   430.61 464.61
- time:provincesb     3   433.89 465.89
- time:lotsize        1   431.16 467.16
<none>                    429.55 467.55
- bank:time           2   445.14 479.14

Step:  AIC=464.1
yield ~ bank + time + lotsize + provincesb + bank:time + bank:lotsize +
    time:lotsize + time:provincesb

                  Df Deviance    AIC
- bank:lotsize     2   433.34 461.34
- time:provincesb  3   435.41 461.41
- time:lotsize     1   432.79 462.79
<none>                 432.10 464.10
- bank:time        2   447.25 475.25

Step:  AIC=461.34
yield ~ bank + time + lotsize + provincesb + bank:time + time:lotsize +
    time:provincesb

                  Df Deviance    AIC
- time:provincesb  3   436.57 458.57
- time:lotsize     1   433.79 459.79
<none>                 433.34 461.34
- bank:time        2   460.03 484.03
```

```
Step:  AIC=458.57
yield ~ bank + time + lotsize + provincesb + bank:time + time:lotsize

               Df Deviance    AIC
- time:lotsize  1   437.21 457.21
<none>              436.57 458.57
- provincesb    3   457.87 473.87
- bank:time     2   463.64 481.64

Step:  AIC=457.21
yield ~ bank + time + lotsize + provincesb + bank:time

             Df Deviance    AIC
<none>            437.21 457.21
- lotsize     1   441.50 459.50
- provincesb  3   458.57 472.57
- bank:time   2   463.82 479.82
There were 13 warnings (use warnings() to see them)
> anova(modl4,test="Chisq")
Analysis of Deviance Table

Model: binomial, link: logit

Response: yield

Terms added sequentially (first to last)


           Df Deviance Resid. Df Resid. Dev P(>|Chi|)
NULL                         690     953.86
bank        2  253.745       688     700.12 < 2.2e-16 ***
time        1  211.606       687     488.51 < 2.2e-16 ***
lotsize     1    3.285       686     485.22    0.0699 .
provincesb  3   21.403       683     463.82 8.682e-05 ***
bank:time   2   26.606       681     437.21 1.669e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> writeLines("\n\n\nTest significance of lotsize term\n\n\n")



Test significance of lotsize term



> modl5<-update(modl4,~.-lotsize)
> anova(modl4,modl5,test="Chi")
Analysis of Deviance Table

Model 1: yield ~ bank + time + lotsize + provincesb + bank:time
Model 2: yield ~ bank + time + provincesb + bank:time
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       681     437.21
2       682     441.50 -1  -4.2811   0.03854 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> writeLines("\n\n\nTest significance of provincesb term\n\n\n")



Test significance of provincesb term



> modl6<-update(modl4,~.-provincesb)
> anova(modl4,modl6,test="Chi")
Analysis of Deviance Table
```

```
Model 1: yield ~ bank + time + lotsize + provincesb + bank:time
Model 2: yield ~ bank + time + lotsize + bank:time
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       681     437.21
2       684     458.57 -3  -21.358  8.87e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> writeLines("\n\n\nTest significance of bank:time interaction\n\n\n")


Test significance of bank:time interaction



> modl7<-update(modl4,~.-bank:time)
> anova(modl4,modl7,test="Chi")
Analysis of Deviance Table

Model 1: yield ~ bank + time + lotsize + provincesb + bank:time
Model 2: yield ~ bank + time + lotsize + provincesb
  Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1       681     437.21
2       683     463.82 -2  -26.606 1.669e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
>
>
>
#***********************************************************************
*********
> #
> #    Confusion Matrix for Simplified Model after excluding third
order term
> #
>
#***********************************************************************
*********
>
> writeLines("\n\n\nConfusion Matrix for Simplified Model p=0.5\n\n\n")


Confusion Matrix for Simplified Model p=0.5



>
> pred<-(rep("Low",length(fitted(modl4))))
> pred<-factor(pred,levels=c("Low","High"))
> pred[fitted(modl4)<0.5]<-"Low"
> pred[fitted(modl4)>=0.5]<-"High"
> table(pred,yield=opdata$yield)
      yield
pred   Low High
  Low  287   13
  High  85  306
>
>
>
#***********************************************************************
***********************
> #
> #     Co-efficients for fitted Simplified Model
> #
```

```
>
#*********************************************************************
************************
> writeLines("\n\n\nCo-efficients for Simplified Model\n\n\n")



Co-efficients for Simplified Model



> coef(modl4)
           (Intercept)             bankHSBC                bankLTSB
            14.2888849          -11.8726021             -13.9236279
              timeLate        lotsizeLarge    provincesbWales & SW
           -20.7362279            0.6118799               0.1297940
     provincesbMidlands provincesbNorth Britain    bankHSBC   :timeLate
             0.4852798            1.7165998              14.0011085
   bankLTSB   :timeLate
            17.6507185
> anova(modl4,test="Chisq")
Analysis of Deviance Table

Model: binomial, link: logit

Response: yield

Terms added sequentially (first to last)


           Df Deviance Resid. Df Resid. Dev P(>|Chi|)
NULL                         690     953.86
bank        2  253.745       688     700.12 < 2.2e-16 ***
time        1  211.606       687     488.51 < 2.2e-16 ***
lotsize     1    3.285       686     485.22    0.0699 .
provincesb  3   21.403       683     463.82 8.682e-05 ***
bank:time   2   26.606       681     437.21 1.669e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> summary(modl4)

Call:
glm(formula = yield ~ bank + time + lotsize + provincesb + bank:time,
    family = binomial, data = opdata)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-3.08332  -0.22025  -0.07639   0.48480   3.41713

Coefficients:
                        Estimate Std. Error z value Pr(>|z|)
(Intercept)              14.2889   697.6201   0.020    0.984
bankHSBC                -11.8726   697.6201  -0.017    0.986
bankLTSB                -13.9236   697.6200  -0.020    0.984
timeLate                -20.7362   697.6207  -0.030    0.976
lotsizeLarge              0.6119     0.3038   2.014    0.044 *
provincesbWales & SW      0.1298     0.3094   0.419    0.675
provincesbMidlands        0.4853     0.3276   1.481    0.139
provincesbNorth Britain   1.7166     0.4232   4.057 4.98e-05 ***
bankHSBC   :timeLate     14.0011   697.6212   0.020    0.984
bankLTSB   :timeLate     17.6507   697.6208   0.025    0.980
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 953.86  on 690  degrees of freedom
Residual deviance: 437.21  on 681  degrees of freedom
AIC: 457.21
```

```
Number of Fisher Scoring iterations: 14

>
>
#***********************************************************************
*************************
> #
> #     Cross Validation of Simplified Model with data divided randomly
into 10 approx.equal groups
> #
>
#***********************************************************************
*************************
>
>
>
>
>
#***********************************************************************
*************************
> #     Randomly divide data into 10 equal groups
>
#***********************************************************************
*************************
> randnos<-with(opdata,runif(length(yield)))
> orderednos<-order(randnos)
> full.length<-with(opdata,full.length<-(length(yield)))
> sub.length<-floor(full.length/10)
> sections<-seq(0,full.length,sub.length)
> sections[11]<-full.length
>
>
>
#***********************************************************************
*************************
> #     Cross Validate
>
#***********************************************************************
*************************
>
> model.formula<-formula(modl4)
>
> cross.val.data<-cbind(opdata,randnos,orderednos)
>
>
> for(i in 1:10){
+ cross.val.sample<-with(cross.val.data,{
+ tempdata0<-cross.val.data[{orderednos<=sections[i] |
orderednos>sections[i+1]},]
+ return(tempdata0)
+ })
+ cross.val.test<-with(cross.val.data,{
+ tempdata1<-cross.val.data[{orderednos>sections[i] &
orderednos<=sections[i+1]},]
+ return(tempdata1)
+ })
+
+ fitted.model<-glm(model.formula,family=binomial,cross.val.sample)
+
+ fitted.values<-
predict(fitted.model,newdata=cross.val.test,type="response")
+ outfile<-cbind(cross.val.test,fitted=fitted.values)
+
+ if(i==1)results<-outfile
+ else results<-rbind(results,outfile)
+
+ }
>
```

```
>
>
>
#***********************************************************************
************************
> #
> #     Confusion Matrix for Simplified Model after Cross Validation
> #
>
#***********************************************************************
************************
> writeLines("\n\n\nConfusion Matrix after Cross Validation -
Simplified Model p=0.5\n\n\n")



Confusion Matrix after Cross Validation - Simplified Model p=0.5



>
> pred<-(rep("Low",length(results$fitted)))
> pred<-factor(pred,levels=c("Low","High"))
> pred[results$fitted<0.5]<-"Low"
> pred[results$fitted>=0.5]<-"High"
> table(pred,yield=results$yield)
      yield
pred   Low High
  Low  287   13
  High  85  306
>
>
>
>
#***********************************************************************
************************
> #
> #     ROC chart for Simplified Model after Cross Validation
> #
>
#***********************************************************************
************************
> writeLines("\n\n\nROC Chart after Cross Validation - Simplified
Model\n\n\n")



ROC Chart after Cross Validation - Simplified Model



>
> simple.pred<-
prediction(results$fitted,results$yield,label.ordering=c("Low","High"))
> simple.perf<-performance(simple.pred,"tpr","fpr")
> plot(simple.perf,main="ROC Chart after Cross Validation - Simplified
Model Provinces B",cex.main=0.9,colorize=TRUE)
> text(locator(1),"Key to Cutoff Probability",srt=90)
>
> perf<-performance(basic.pred,"prec","fpr")
> plot(perf,main="Precision versus False Positive Error Rate Provinces
B",cex.main=0.9,
+            sub="Simplified Model",cex.sub=0.9,
+            colorize=T)
> text(locator(1),"Key to Cutoff Probability",srt=90)
>
>
```

```
>
#*************************************************************************
***********************
> #     Output ROC chart data for Simplified Model after Cross
Validation
>
#*************************************************************************
***********************
> rocout<-data.frame(fitted=results$fitted,yield=results$yield)
> write.table(rocout,file="provbsimpleroc.out")
>
>
>
#*************************************************************************
*******
> #
> #                              bar chart
> #
>
#*************************************************************************
*******
>
> limits<-seq(0.0,1.0,0.1)
> temppred<-rep(1,length(results$fitted))
>
> predgrp<-with(results,{
+ for(i in 1:10)temppred[fitted>limits[i] & fitted<=limits[i+1]]<-i
+ return(temppred)
+ }
+ )
>
> dummy<-rep(1,length(predgrp))
> add.results<-cbind(results,pred.grp=predgrp,counter=dummy)
>
>
> #"barchart.data" is the array containing the number of low yield and
high yield observations
> # whose prdicted probabilites of being in thehigh yield group are in
the ranges
> #specified by "intervals"
>
> barchart.data<-
with(add.results,tapply(counter,list(yield,pred.grp),sum))
>
> intervals<-c("0.00-0.09","0.10-0.19","0.20-0.29","0.30-0.39","0.40-
0.49",
+  "0.50-0.59","0.60-0.69","0.70-0.79","0.80-0.89","0.90-1.00")
>
>
barplot(barchart.data,beside=TRUE,names.arg=intervals,space=c(1,5),cex.
names=0.5,
+   main="Counts against Probability(high yield) Provinces
B",cex.main=0.9,
+         sub="Simplified Model",cex.sub=0.9,
+   xlab="Probability case belongs to high yield group",
+   ylab="No of cases",
+   col=c("grey25","grey75"))
>
> legend(locator(1),legend=c("low yield","high
yield"),fill=c("grey25","grey75"))
>
>
>
>
#*************************************************************************
*******
> #
> #         ROC Chart comparing Basic and Simplified Models
> #
```

```
>
#*********************************************************************
*******
> writeLines("\n\n\nROC Chart comparing Basic and Simplified
Models\n\n\n")



ROC Chart comparing Basic and Simplified Models



>
>
> basic.perf<-performance(basic.pred,"tpr","fpr")
> plot(basic.perf,col="blue",main="ROC Chart comparing Basic and
Simplified Models Provinces B",cex.main=0.9)
> plot(simple.perf,col="red",add=TRUE)
> legend(locator(1),legend=c("basic model","simplified
model"),fill=c("blue","red"))
>
>
>
#*********************************************************************
*******
> #
> #                remove following "#" if outputing to file
> #
>
#*********************************************************************
*******
>
> #sink(file=NULL)
>
>
>
>
```

Appendix IX

Provinces B ANOVA plots

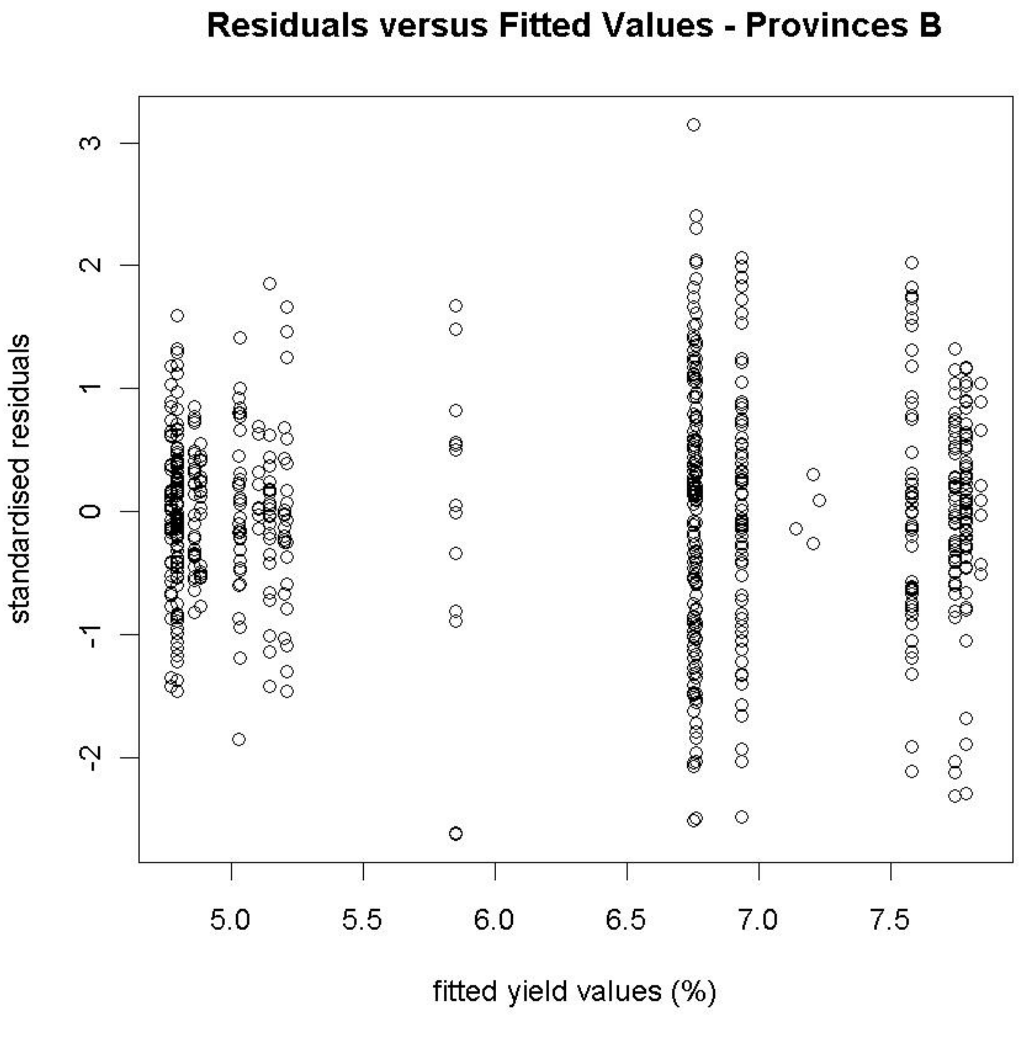


Figure IX.1 Residuals versus Fitted Values – *Provinces B*

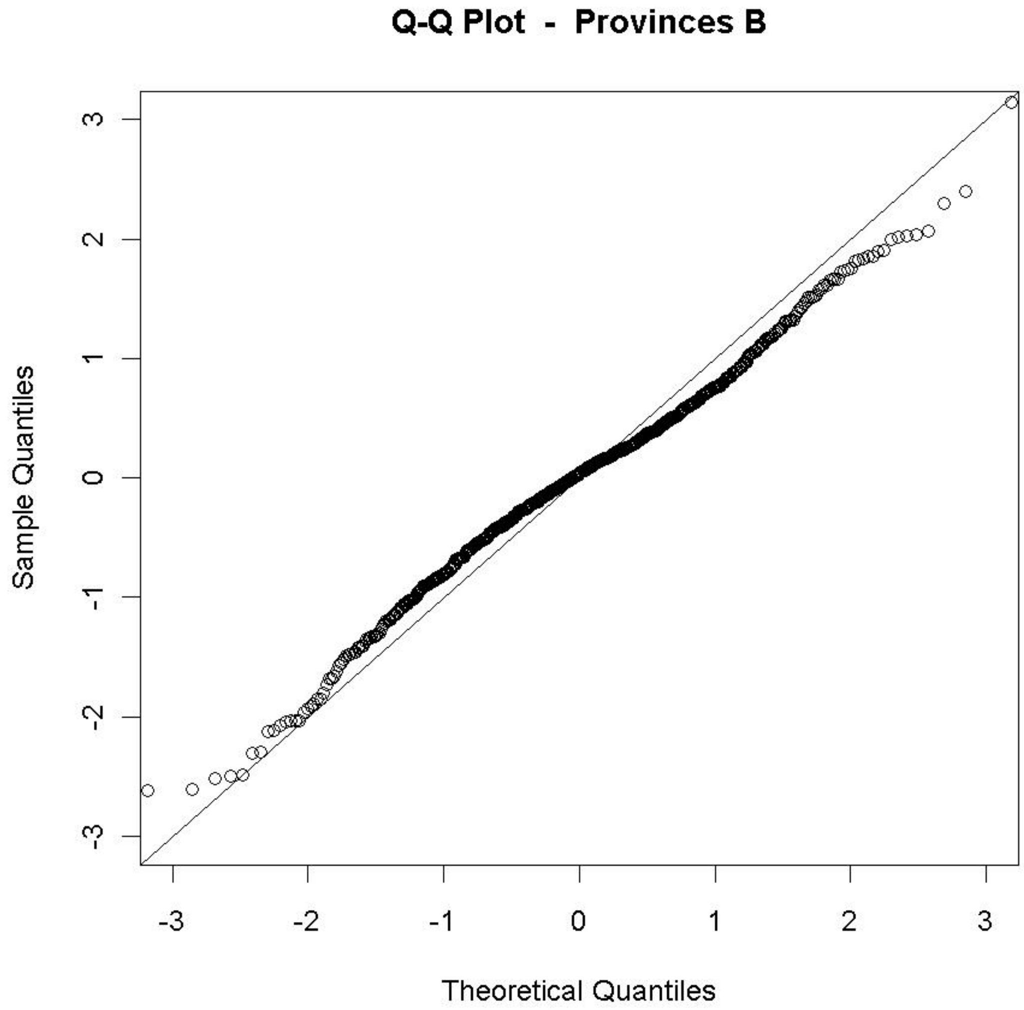


Figure IX.2 Q-Q Plot – Provinces B

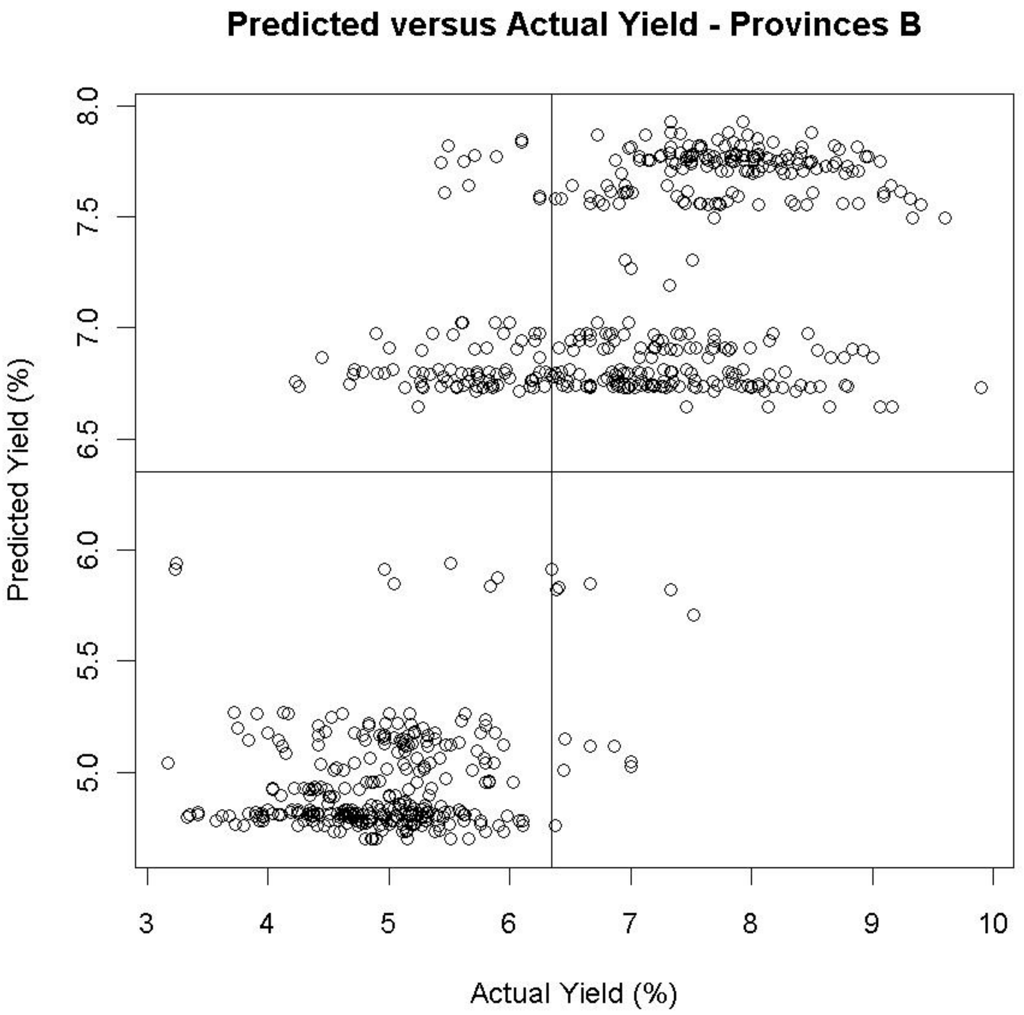


Figure IX.3 Predicted versus Actual Yield – *Provinces B*

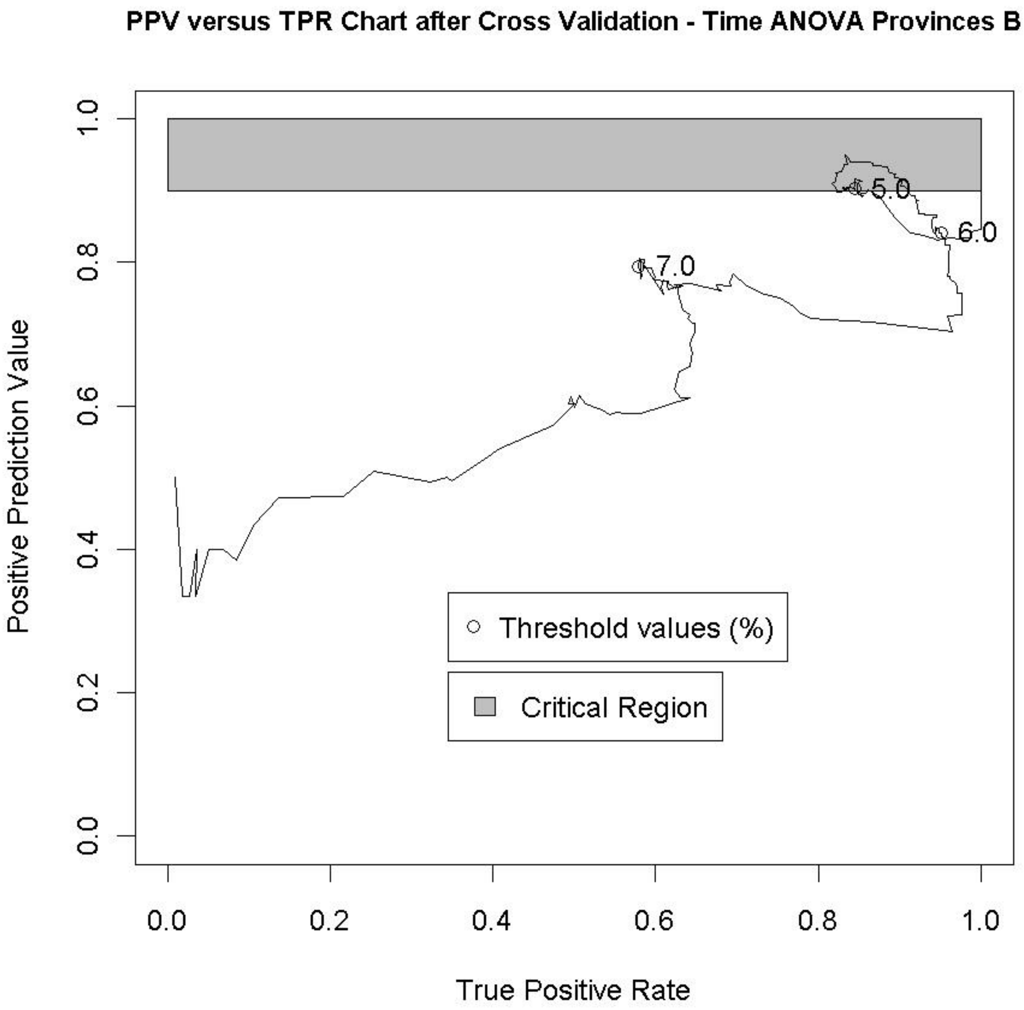


Figure IX.4 PPV versus TPR Chart after Cross-Validation – Time ANOVA *Provinces B*

Appendix X

Post-study dataset (October, 2006 – December, 2007)

Case No	Date	Bank	Region	Location	Price £	Rent £	Yield (%)	Lot Size	Tenure	Provinces A
1	18 Oct 2006	LTSB	East Midlands	NORTHAMPTON	1150000	74000	6.43	L	F	Midlands
2	18 Oct 2006	LTSB	South-East	AYLESBURY	2230000	90000	4.04	L	F	South
3	18 Oct 2006	LTSB	West Midlands	COVENTRY High St	2610000	139500	5.34	L	F	Midlands
4	18 Oct 2006	LTSB	South-East	SOUTHAMPTON	2210000	115000	5.20	L	F	South
5	18 Oct 2006	HSBC	East Anglia	TRING	570000	28000	4.91	L	F	Midlands
6	18 Oct 2006	HSBC	South-East	EASTLEIGH	805000	35000	4.35	L	F	South
7	18 Oct 2006	HSBC	West Midlands	BLOXWICH	1075000	62500	5.81	L	F	Midlands
8	18 Oct 2006	Barclay	London-M25	RAINHAM	995000	42460	4.27	L	F	South
9	18 Oct 2006	HSBC	South-East	HORSHAM	1705000	75000	4.40	L	F	South
10	18 Oct 2006	LTSB	East Midlands	NOTTINGHAM	3525000	170000	4.82	L	F	Midlands
11	18 Oct 2006	LTSB	Scotland	GALASHIELS	670000	40000	5.97	L	Feuhold	Celt
12	18 Oct 2006	HSBC	East Midlands	SWADLINCOTE	950000	45000	4.74	L	F	Midlands
13	18 Oct 2006	HSBC	North-East	CHAPEL ALLERTON	940000	40000	4.26	L	F	North
14	18 Oct 2006	LTSB	North-East	KINGSTON HULL Silver	750000	39000	5.20	L	F	North
15	18 Oct 2006	HSBC	North-East	Hedon	455000	21000	4.62	M	F	North
16	06 Dec 2006	Barclay	North-West	SALE	650000	29000	4.46	L	F	North
17	06 Dec 2006	Barclay	North-West	CHESTER	500000	23000	4.60	M	F	North
18	06 Dec 2006	Barclay	South-East	CRANBROOK	392500	14900	3.80	M	F	South
19	06 Dec 2006	Barclay	South-West	WOOTTON BASSETT	455000	20850	4.58	M	F	South
20	06 Dec 2006	Barclay	East Anglia	HAVERHILL	1100000	53500	4.86	L	F	Midlands
21	06 Dec 2006	Barclay	North-East	GRIMSBY	1790000	95000	5.31	L	F	North
22	06 Dec 2006	Barclay	North-East	DURHAM New Elvet	965000	37000	3.83	L	F	North
23	06 Dec 2006	Barclay	North-East	HETTON-LE-HOLE	505000	25000	4.95	L	F	North
24	06 Dec 2006	Barclay	North-East	GOSFORTH	730000	28000	3.84	L	F	North
25	06 Dec 2006	LTSB	South-West	ST AUSTELL	660000	37000	5.61	L	F	South
26	06 Dec 2006	Barclay	East Anglia	HALESWORTH	455000	20800	4.57	M	F	Midlands
27	06 Dec 2006	Barclay	South-East	WOKINGHAM	1600000	78000	4.88	L	F	South
28	06 Dec 2006	Barclay	South-East	BRIGHTON Preston R	750000	28500	3.80	L	F	South
29	06 Dec 2006	LTSB	Wales	BRYNMAWR	210000	9000	4.29	S	F	Celt

Appendix X (Continued)

Post-study dataset (October, 2006 – December, 2007)

Case No	Date	Bank	Region	Location	Price	Rent	Yield	Lot Size	Tenure	Provinces A
30	06 Feb 2007	Barclay	London-M25	ACTON	1010000	42600	4.22	L	F	South
31	06 Feb 2007	Barclay	South-West	SHERBORNE	570000	20800	3.65	L	F	South
32	06 Feb 2007	Barclay	West Midlands	BEWDLEY	395000	17500	4.43	M	F	Midlands
33	06 Feb 2007	Barclay	South-East	SUTTON	490000	19900	4.06	M	F	South
34	06 Feb 2007	HSBC	London-M25	DEPTFORD	620000	30500	4.92	L	F	South
35	06 Feb 2007	Barclay	West Midlands	KIDSGROVE	325000	16000	4.92	M	F	Midlands
36	06 Feb 2007	Barclay	East Midlands	THRAPSTON	452500	20000	4.42	M	F	Midlands
37	06 Feb 2007	Barclay	East Anglia	SAWSTON	640000	27000	4.22	L	F	Midlands
38	06 Feb 2007	Barclay	South-East	HUNGERFORD	390000	18600	4.77	M	F	South
39	06 Feb 2007	Barclay	West Midlands	SHELDON	730000	35000	4.79	L	F	Midlands
40	06 Feb 2007	Barclay	North-East	PUDSEY	402000	20000	4.98	M	F	North
41	06 Feb 2007	Barclay	North-East	LOW FELL	375000	17500	4.67	M	F	North
42	06 Feb 2007	Barclay	South-East	HORSHAM	1730000	63300	3.66	L	F	South
43	06 Feb 2007	Barclay	North-West	ALDERLEY EDGE	1010000	47500	4.70	L	F	North
44	06 Feb 2007	LTSB	South-East	MARLOW	4030000	147000	3.65	L	F	South
45	06 Feb 2007	Barclay	South-West	CHELTENHAM	1930000	85000	4.40	L	F	South
46	06 Feb 2007	Barclay	North-West	KENDAL	1690000	87400	5.17	L	F	North
47	06 Feb 2007	Barclay	West Midlands	COVENTRY Walsgrove	410000	17000	4.15	M	F	Midlands
48	06 Feb 2007	Barclay	East Midlands	GAINSBOROUGH	460000	20000	4.35	M	F	Midlands
49	06 Feb 2007	Barclay	North-East	KINGSTON HULL Hold	535000	26000	4.86	L	F	North
50	06 Feb 2007	Barclay	North-West	STALEYBRIDGE	400000	16500	4.13	M	F	North
51	06 Feb 2007	Barclay	North-West	WHALLEY	430000	19500	4.53	M	F	North
52	06 Feb 2007	HSBC	North-East	HECKMONDWIKE	450000	22000	4.89	M	F	North
53	26 Mar 2007	Barclay	South-East	SAWBRIDGEWORTH	545000	25000	4.59	L	F	South
54	26 Mar 2007	Barclay	South-East	OLNEY	470000	20366	4.33	M	F	South
55	26 Mar 2007	Barclay	South-East	RAINHAM Kent	790000	40000	5.06	L	F	South
56	26 Mar 2007	Barclay	North-West	ORMSKIRK	1000000	49250	4.92	L	F	North
57	26 Mar 2007	Barclay	South-West	CULLOMPTON	655000	30000	4.58	L	F	South
58	26 Mar 2007	HSBC	West Midlands	NEWPORT	340000	14000	4.12	M	F	Midlands

Appendix X (Continued)

Post-study dataset (October, 2006 – December, 2007)

Case No	Date	Bank	Region	Location	Price	Rent	Yield	Lot Size	Tenure	Provinces A
59	26 Mar 2007	Barclay	North-West	BROMBOROUGH	405000	20000	4.94	M	F	North
60	26 Mar 2007	Barclay	London-M25	ROTHERHITHE	660000	25200	3.82	L	F	South
61	26 Mar 2007	Barclay	London-M25	SE23 0AX	1315000	51000	3.88	L	F	South
62	26 Mar 2007	Barclay	North-East	LEEDS LS2 9HG	810000	29750	3.67	L	F	North
63	26 Mar 2007	Barclay	London-M25	RAINHAM RM13	855000	42460	4.97	L	F	South
64	26 Mar 2007	Barclay	West Midlands	WALSALL	1260000	65000	5.16	L	F	Midlands
65	21 May 2007	Barclay	South-East	ABBOTTS LANGLEY	410000	19000	4.63	M	F	South
66	21 May 2007	Barclay	East Midlands	ASHBOURNE	410000	20000	4.88	M	F	Midlands
67	21 May 2007	Barclay	East Anglia	Histon CB24	380000	16000	4.21	M	F	Midlands
68	21 May 2007	Barclay	West Midlands	COVENTRY	615000	26500	4.31	L	F	Midlands
69	21 May 2007	Barclay	North-West	WORKINGTON	500000	23600	4.72	M	F	North
70	21 May 2007	Barclay	South-West	WARMINSTER	640000	28500	4.45	L	F	South
71	21 May 2007	HSBC	London-M25	ILFORD	2110000	120000	5.69	L	F	South
72	21 May 2007	Barclay	South-East	SUTTON	1390000	78600	5.65	L	F	South
73	21 May 2007	Barclay	North-East	HOUGHTON-LE-S	430000	23000	5.35	M	F	North
74	21 May 2007	Barclay	East Anglia	LEISTON	375000	18000	4.80	M	F	Midlands
75	21 May 2007	Barclay	North-East	YORK YO24 4LZ	490000	25000	5.10	M	F	North
76	10 Jul 2007	HSBC	South-West	TROWBRIDGE	1100000	60000	5.45	L	F	South
77	10 Jul 2007	Barclay	London-M25	RAINHAM RM13	795000	42460	5.34	L	F	South
78	10 Jul 2007	Barclay	West Midlands	LONGTON	410000	20000	4.88	M	F	Midlands
79	10 Jul 2007	Barclay	North-East	GATESHEAD	880000	45000	5.11	L	F	North
80	10 Jul 2007	HSBC	East Anglia	WOODBRIDGE	740000	40000	5.41	L	F	Midlands
81	10 Jul 2007	HSBC	South-West	FALMOUTH	1680000	88500	5.27	L	F	South
82	10 Jul 2007	HSBC	Wales	ABERYSWYTH	1400000	85000	6.07	L	F	Celt
83	10 Jul 2007	HSBC	North-West	BLACKPOOL	1050000	60000	5.71	L	F	North
84	10 Jul 2007	HSBC	North-West	PENRITH	840000	42500	5.06	L	F	North
85	11 Oct 2007	HSBC	East Midlands	OAKHAM	650000	40000	6.15	L	F	Midlands
86	11 Oct 2007	Barclay	West Midlands	WILLENHALL	320000	12500	3.91	M	F	Midlands
87	11 Oct 2007	Barclay	East Midlands	NOTTINGHAM 2QP	302500	12500	4.13	M	F	Midlands
88	11 Oct 2007	HSBC	London-M25	WALTHAM CROSS	720000	36000	5.00	L	F	South
89	11 Oct 2007	Barclay	London-M25	ENFIELD 0BZ	615000	23000	3.74	L	F	South

Appendix X (Continued)

Post-study dataset (October, 2006 – December, 2007)

Case No	Date	Bank	Region	Location	Price	Rent	Yield	Lot Size	Tenure	Provinces A
90	11 Oct 2007	HSBC	North-East	WHITLEY BAY	525000	30000	5.71	L	F	North
91	11 Oct 2007	HSBC	North-East	ALNWICK	535000	30000	5.61	L	F	North
92	11 Oct 2007	Barclay	North-West	BIRKENHEAD	407500	21000	5.15	M	F	North
93	11 Oct 2007	Barclay	North-West	HYDE	505000	30000	5.94	L	F	North
94	11 Oct 2007	HSBC	East Anglia	SUDBURY Suffolk	865000	40000	4.62	L	F	Midlands
95	11 Oct 2007	HSBC	East Anglia	ST IVES Cambs	555000	32000	5.77	L	F	Midlands
96	11 Oct 2007	Barclay	East Anglia	DOVERCOURT	420000	16000	3.81	M	F	Midlands
97	11 Oct 2007	Barclay	North-East	KINGSTON HULL 7RB	455000	24000	5.27	M	F	North
98	11 Oct 2007	HSBC	North-East	SOUTH SHIELDS	515000	34000	6.60	L	F	North
99	11 Oct 2007	HSBC	East Anglia	NORTH WALSHAM	640000	30500	4.77	L	F	Midlands
100	11 Oct 2007	HSBC	West Midlands	DUDLEY	770000	47500	6.17	L	F	Midlands
101	11 Oct 2007	Barclay	North-West	LIVERPOOL 4RA	250000	12850	5.14	S	F	North
102	01 Dec 2007	Barclay	London-M25	E3 3AB	550000	27000	4.91	L	F	South
103	01 Dec 2007	LTSB	South-East	BATTLE	342500	15700	4.58	M	F	South
104	01 Dec 2007	Barclay	North-West	LYMM	560000	31000	5.54	L	F	North
105	01 Dec 2007	Barclay	South-East	HARPENDEN	1040000	51500	4.95	L	F	South
106	01 Dec 2007	LTSB	South-East	EAST GRINSTEAD	1470000	64000	4.35	L	F	South
107	01 Dec 2007	LTSB	West Midlands	WOLVERHAMPTON	4300000	245000	5.70	L	F	Midlands
108	01 Dec 2007	Barclay	South-East	BROMLEY Kent	2900000	159000	5.48	L	F	South
109	01 Dec 2007	LTSB	North-West	MORECAMBE	655000	41500	6.35	L	F	North
110	01 Dec 2007	Barclay	East Midlands	LENTON Notts	565000	24500	4.34	L	F	Midlands

Appendix XI

Published Works

Table XI.1 Papers published during the study

Title of Paper	Place of Publication	Date of Publication	Author/s	Peer Reviewed
The Impact of the Time Value of Money on Valuation Practice.	FIG 2006 Congress, Munich.	October, 2006.	Malvern Tipping	Yes
Sale-and-leaseback as a British Real Estate Model.	FIG 2007 Congress, Hong Kong.	May, 2007.	Malvern Tipping and Richard Bullard	Yes
Sale-and-leaseback as a British Real Estate Model.	Journal of Corporate Real Estate.	December, 2007.	Malvern Tipping and Richard Bullard	Yes, double-blind peer reviewed
Identifying Clay-construction Buildings in a Norfolk Market Town.	FIG 2010 Congress, Sydney.	April, 2010.	Malvern Tipping	No
Factors Most Influencing United Kingdom Retail Property Investment Yields: A Theoretical Perspective.	COBRA 2010, Paris.	September, 2010.	Malvern Tipping and Terence Lam	Yes