QUANTIFICATION OF THE RISK ASSOCIATED WITH THE SEASONAL FINANCING OF AGRICULTURAL PRODUCTION

A thesis presented in partial
fulfilment of the requirements
for the degree of

Master of Agricultural Economics
Massey University

Ron Veltman

January 1992

ABSTRACT

Since the abolition of government support policies for both agricultural and financial industries during the early 1980s, participants have had to take direct responsibility for the management of the risks involved in their business activity. As a prerequisite to the development of practical risk management strategies and techniques, quantification of risk is considered by this thesis.

A quantification risk index that incorporates both the third and fourth moments of a distribution, thus adding to variance and monotonic transformations, the traditional surrogate risk measures, was developed and applied to sheep and beef farming.

The risk index is developed using logit analysis, where risk is directly estimated. Logit analysis was used because it suited the thesis definition of risk. In this thesis, risk is defined as the probability of incurring loss or harm, where loss or harm is defined, in the context of sheep and beef farming, as zero or less than zero 'net cash returns'. Net cash returns are defined as all cash revenues generated by farm production less all farm and farmer expenditures. The index, or probability, is directly estimated given forecast average market prices, effective farm area, total farmer forecast expenditures and island location (North or South).

The risk index has been developed for banker application to farm budgets submitted for the purposes of seasonal finance approval. The banker is warned by the index that the proposed farm plan has a high probability of ending in farm insolvency and an inability of the farmer to service all lending in the forthcoming year, solely from farm production.

As a consequence of applying the measure to sheep and beef farming, the thesis found that in terms of risk to net cash returns, effective farm area in conjunction with total farmer expenditure is significantly ranked higher than fluctuating market product prices, and that risk trade-offs exist between farm area and expenditures. In a situation of small farm size with relatively high expenditures, optimistic product prices are insufficient to offset the high probability of incurring negative net cash returns.

ACKNOWLEDGEMENTS

I wish to acknowledge the help and support of both Professor Alan Rae and Dr Jeff Weber in their constructive supervision of this thesis. I also thank my wife and two children for their understanding and endurance of my frequent abstinence from normal family life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter One

INTRODUCTION

1.1 THE AGRICULTURAL SECTOR AND RISK

1.1.1 BACKGROUND

New Zealand agriculture has undergone considerable change in direction over the past thirty years. After the relative stability but slow general economic growth of the 1960s, the 1970s saw the evolution of a combination of government policies designed to initially stimulate economic growth and later, after an OPEC driven redistribution of the world's income in its favour, protect export agricultural production from the sudden consequential market contraction of industrialised and oil importing countries, and the associated long term price decline in the international commodity market (Hawke 1987).

Reaction to depressed international prices for agricultural commodities included the supplementation of dwindling producer incomes, a policy that was in addition to a growing list of regulatory and interventionist measures designed to compensate agriculture for the costs of import protection and of maintaining an over-valued exchange rate (Hawke 1987).

By 1984 the interventions in place consisted of direct input subsidies (fertiliser, irrigation, interest), production subsidies (supplementary minimum prices), development schemes (livestock incentive scheme, land development encouragement loan), the provision of research and farm services, producer board subsidies, taxation exemptions, industry controls and producer board legislation, as well as state ownership of the rural banking and finance corporation. The 1984 total fiscal cost of these interventions were estimated to be 3.2 percent of gross domestic product (GDP), or $1,087 million. (Raynor 1987).

Under this 'protectionist' environment, producers and agri-service industries protected from the business risks associated with international trade had little incentive to adapt to a changing international commodity market, let alone implement systems of business risk management at the micro level. The early 1980s saw the recognition that the international downturn was not short term, resource utilisation inefficiencies had developed within the economy, overseas borrowing was not sustainable and those in agri-business at all levels had to directly confront the realities of the international marketplace if they were to quickly adjust to the international environment (Hawke 1987).

As a consequence, protectionist policies were removed by the new Labour government elected in 1984. The result has been the removal of those many anomalies that encouraged inefficient use of resources (Pryde,Bain 1985). Since that period the majority of businesses involved in agricultural activity have had to directly face the undisguised risks associated with their involvement. Given that risk is a predominant feature of all agricultural activity, and agri-business must now take direct responsibility for risk, there is now demand for practical systems of risk evaluation and management.

1.1.2 AGRICULTURAL RISK

Within portfolio theory, risk is seen as being comprised of both systematic and nonsystematic elements, i.e., nondiversifiable and diversifiable. This concept tries to divide risk into those risk components that are inherent in all agricultural activity (systematic), and those risk elements that are able to be eliminated through diversification into other investment options. Turvey and Driver (1986) in a study designed to determine the extent of systematic and nonsystematic risk within United States agriculture, concluded that there is in fact a great deal of systematic risk in agriculture and the proportion of local and specific risk that can be diversified away is small, relative to the total risk of the farm sector portfolio.

Opportunities for diversification from one generic type of agriculture into a better alternative are limited by the constraints imposed by all manner of resource factors. For instance, a change from Merino sheep farming to intensive horticulture may not be feasible due to the absence of those basic physical resources, required by horticulture, such as soil type or appropriate climate.

Given the range of physical resources that are available to any specific farmer, the product option range within generic types of agriculture is also limited by their comparative suitability to those resources. For example Romney sheep would be less suited to South Island high country conditions than Merino sheep.

Once a farmer is committed to a production decision, his options diminish as he draws nearer to harvest. For instance, an arable crop farmer may harvest peas in either their green state, as fresh peas, or in their dry seed state. Once the time has passed for green pea harvest, the farmer is committed to the dry pea market.

Agricultural risk and uncertainty can be further divided into business or financial risk, where business risk can be further categorised according to three sources (Just 1975). First, risk can be associated with environmental factors such as climate, disease, pest infestation and technological obsolescence. Second, risk can arise from market factors such as supply and demand disequilibrium within both input and output markets, and competitive elements associated with market structure. Third, risk can occur from government policy and programs, such as support levels and regulations, as well as government priorities with regard to economic objectives.

During the season of production, from investment to harvest, a farmer must not only endure the risks and uncertainties associated with any probable change in, or occurrence of, any of the above risk sources, but also the risks associated with their resulting impact or outcome.

The risk incurred as a consequence of a change in any of the above risk elements within the three source categories can vary across the national agricultural economy. Each risk source has a regional qualification. For example, the probability, or risk, of drought on the East Coast is greater than on the West; the risk associated with local price fluctuations for market garden produce are less in the Auckland region than in the Invercargil region.

Risk can also vary according to the 'additive' or cumulative combination of separate risk sources within regions. For example, the perceived risk to, or impact on, cashflows as a consequence of the 1984 SMP removal phase of government policy would have been greater in the Canterbury region than in the West Coast region because Canterbury is much more prone to prolonged drought, meaning that cashflows were already at risk prior to SMP removal. The implication is that the level of risk associated with the occurrence of any specific stimuli on any generic agricultural activity in any particular region differs from other regions according to the combined probability of risk stimuli occurring concurrently at any point in time.

The level of risk might be determined by the impact of any stimuli rather than the probable occurrence of stimuli. More specifically, the concept of risk might be more appropriately associated with the probable impact of say a large price decrease rather than the occurrence of such a decrease. Total risk within a farm unit describes the combined probable impact upon the security of the farm unit, of the occurrence of any combination of stimuli.

Amongst those industries that service agricultural production, banking is the industry that is an important prerequisite agri-service input common to all agricultural and aquacultural production and associated activities. This fact makes agricultural production dependent on the security and risk exposure perceptions that a bank may have with regard to agricultural finance involvement. A bank is in a position to determine the productive longevity of any individual producer, using its ability to either invest or disinvest in its farmer client.

As a consequence, a feature of agricultural risk, from a producer's point of view, is the financier's reaction to the impact upon the producer of any adverse change in any risky factor. This form of risk can be categorised as financial risk, and is best described by the following scenario.

Assume that during October a sheep farmer has entered an unexpected drought period. Further, his June negotiated seasonal overdraft facility requires that he receives better than average prices for his produce and his expenditure be at a controlled minimum level. The farmer's overdraft includes a drought allowance that facilitates the purchase of a small quantity of stock feed should it be required. Overall this farmer's debt equity ratio is dangerously large but his current account deficit not unusually large for this time of the year. He is classified by the bank as being a 'security borderline' client.

As the drought progresses into January, it becomes apparent that the stock feed allowance will need to be spent, and an 'insufficient' quantity of feed barley is purchased due to the unit price paid being beyond budget expectation as a consequence of the high regional demand for feed barley, generated by the drought. Further, he sells a large proportion of his prime lambs earlier than expected, to ease the immediate stock demand for pasture, at prices less than budget expectations.

The farmer believes he has reacted sensibly to his situation, in that he is implementing decisions designed to protect his future production from the impact of the drought, thus minimising his long-term loss. His concern centres around controlling the weight loss of his ewes prior to mating and shearing, as well as maintaining a small weight gain in his replacement ewe lambs. He is also concerned with reducing his short term loss by selling lambs before the market further worsens.

As a consequence of his early sale of lambs and delayed purchase of stock feed, his overdraft facility is now certain of being exceeded. At this point he visits his bank manager with an application to extend his current account overdraft facility. He believes that he was justified in selling his lambs early, as they would not have reached budget target liveweights and grades under the circumstances, and prices offered by the meat companies were worsening as a consequence of the unusually high regional supply of lambs from farmers concurrently wishing to sell lambs early.

Although it is certain that the farmer will sustain an accounting loss and show a consequential deterioration in his equity position, if the bank is prepared to accommodate the overdraft extension then, from the farmer's point of view, the impact of the drought will not be so bad. The farmer has no idea how the bank will react to his application, and considers a possible adverse bank reaction to his plight as an additional risk component within the total risk of drought and its final impact. Should the bank not accept his reasoning and 'harden' their position, then the final impact of the drought on this farmer would be particularly harmful.

The above scenario serves to illustrate two points. First it conveys the concept that risk in agriculture is not only related to the impact on cash outcomes of various uncontrollable stimuli, whatever they may be, but also that the severity, or level of risk again differs according to the impact that the resulting cash outcome may have on the attitudes and business decisions of those financing agriculture, and their consequential influence on the ability of the farmer to continue farming. Second, financiers of agricultural production, once committed to a level of financial involvement, either run the risk of inadvertently underwriting those components of agricultural business risk they feel uncomfortable with, or be requested to do so, thus placing them in the unenviable position of having to decide whether or not to exert great pressure on their client.

Agricultural risk is multi-dimensional in terms of its wide range of source stimuli, and the variability across farmers and regions of the impact that those stimuli may have on the farm unit. Although sources of risk can be categorised and the probability of those source stimuli occurring are known, the quantifiable level of risk itself, or the impact of those stimuli, is not known. Risk can only be described according to the consequential harmful impact of the occurrence of any uncontrollable agricultural characteristic.

1.2 THE FINANCIAL SECTOR AND RISK

1.2.1 BACKGROUND

At the time the Labour party came to government office during 1984, New Zealand's financial sector was the most regulated in the Western world. Since 1984, the financial sector has been completely overhauled. The removal of blanket regulations controlling lending and deposit rates and the abolition of the penal marginal ratios to financial institutions preceded the revocation of both the 30 day rule, which had prevented the trading and savings banks from entering the short end of the money market, and the 3 percent interest rate restriction on the ordinary accounts of savings banks. The removal of credit growth guidelines, foreign exchange controls and the liberalisation of bank registration were also introduced to increase competition and efficiency in the financial system. With the ability to borrow or lend offshore, and substantial changes to the system of tendering for Treasury bills, banks are now in a position to openly compete for custom (Russell 1985).

Deregulation has seen an increase in the number of banks from the major four in 1984 to over 20 in 1990. A growing trend is the number of mergers into large supermarket type banks offering a wide range of financial services. Smaller banks are emerging as niche banks, filling the gaps left by major

1.2.2 THE ROLE OF THE TRADING BANK

Whilst it is difficult to be precise in defining a 'bank' or banking business, a financial institution is part of the banking system if its main functions include the acceptance of demand deposits, the operation of money transfer, and the creation of demand deposits through the making of loans and provision of overdraft credit (Deane 1982).

It is the ability of trading and commercial banks to create money in the form of demand deposits by making loans and extending credit, that distinguishes them from other financial institutions. The creation of deposits can continue so long as banks hold sufficient currency and reserves to meet regulatory requirements and to redeem whatever amounts the holders of deposits want to convert to currency (Crosse 1979). The money a bank can lend or invest, at any point in time, is its excess of cash and bank balances over required reserves and minimum cash requirements, according to its daily balance sheet. The bank must stand ready to pay out the deposits it creates when it makes new loans or extends overdraft facilities.

The creation of demand deposits through overdraft extension is particularly suited to the characteristics of agricultural production. By supplying liquidity to producers, through their ability to lend and invest, they are able to provide money, at a cost, in consideration of assets or effort that have a future money value.

Given that the creation of deposits through lending is directly related to their depository function, i.e., the demand deposits that constitute the major portion of the money supply, then their ability to create deposits is constrained by not only the willingness of customers to deposit funds, but also by the total pool of funds available for deposit.

Where the pool of funds available for deposit is itself constrained and a major proportion of those deposits is transacted for consumption, i.e., they are short duration deposits, then the bank has the essential role of apportioning or rationing the available long term deposits, as credit, to what it perceives as being the most efficient users of that credit.

1.2.3 RISK IN BANKING

Risk within the banking industry can be categorised according to four basic sources. First, market risk broadly consists of elements such as the general state of the economy and competition within the banking industry. Second, political risk includes that risk inherent in changes of government policy as well as the internal management politics often found within large corporations and their boards of directors. Both of these categories can be included within the broad category of business risk.

The third category, deposit risk, identifies the risks associated with the liability side of a bank's balance sheet. Demand deposits present risk according to both the term and size of the deposit. On call deposits present the greatest risk within this category. Deposit risk is included within financial risk, along with the fourth category, credit or default risk, which describes those risks associated with the asset side of a bank's balance sheet. The non-payment of either a loan principal at maturity or interest at any stage during the term of a loan are the main sources of this type of risk.

Default risk is the main emphasis of this thesis and, as the title suggests, concentrates on the short term seasonal provision of working capital to agricultural producers.

Risk is a banker's preoccupation. If loans are not repaid then the banker in turn will not be able to meet his commitments. In this way both deposit and default risk are linked. Risk is inherent in the choice of borrower; risk is implicit in the industry being financed; risk by the business to which the banker may grant too much or too little credit; risk which involves the whole economy; a gambler's risk with weather, geography, technology and politics (Camu 1977).

In making innumerable loans to thousands of undertakings of all sizes, large banks are able to protect themselves against at least the consequences of risk, if not against risk itself. The banker covers himself against risk by guarantees and securities with solid legal backing (Camu 1977).

A bank manager faced with a request for loan or overdraft facility is generally concerned with the answers to four basic questions:

(i) How much does the customer want to borrow?
(ii) What does he want it for?
(iii) How long does he want it for?
(iv) How is it to be repaid?

The four questions are all related to the security aspects of minimising default risk. Although security is not often directly questioned, it underlies the reason for asking these questions in the first place (Cox 1979).

The 'how much' question establishes not only the ability of the bank to feasibly provide such amount according to its current balance of excess demand deposits, but also to ascertain the ratio between what the customer himself is providing against what the bank is being asked to provide. Bank policy generally determines a maximum provision ratio according to the realisable value of the asset being financed, and ensures some equity to the borrower thus guaranteeing a collateral for the loan (Cox 1979).

For short term finance, as well as long term finance, the 'how much' question relates to the 'what for' question. This question is related to the credit use issue where the manager is attempting to establish the security parameters of his involvement. Is the activity high risk, i.e., is it speculative? Will the activity generate a sustainable interest yield? Will the amount requested hinder the ability of the operation to preform to expectation in terms of loan servicing? Are the characteristics of the operation such that it is vulnerable to a whole range of uncontrollable influences? Is the customer sufficiently knowledgable of the operation?

This question also begins to address the security question directly. What assets will be used as collateral for the loan? Do these assets have a realisable value, and how easy are these assets 'cashed in'? For seasonal finance, future production is sufficient collateral if the value of future production exceeds the value of the seasonal finance (Cox 1979).

Because the bank's current liability constitutes short term notice of demand deposit payment, it makes good sense to have loans out on a short term basis. The 'how long' question relates the term and type of loan to the nature of its intended use.

The 'how is it to be repaid' question is related to how the loan will provide for future operation earnings. Will the nature of the operation provide sufficient future profit from which the loan can be repaid? Will repayments be such that they ensure regular cash flow to the bank, i.e., will payments be monthly etc.? Will the loan realise a regular yield, or will the return from the loan occur at some future date? (Cox 1979).

Apart from answers to the above questions, other factors play a vital part in linking together the answers to the four basic questions when evaluating the loan application. One highly variable factor is the customer. A bank manager must get to know his customer's health, age, activity and the value of connected family and business accounts held at the branch. This is vital information to the manager (Cox 1979). It serves to not only indicate to the bank manager any possible sources of default risk, but also the extent of possible guarantees and securities that are at his disposal. His objective is to increase bank assets by lending to earn interest revenue and in so doing, help the customer by providing a loan with such security that the risks are minimised for both parties (Cox 1979).

1.3 THE THESIS DEFINITION OF RISK

1.3.1 RISK DEFINED

Risk is not an observable entity - it is a concept verbally defined by Websters dictionary as 'the chance of injury, damage or loss'. Although the verbal definition is intuitively appealing, it would not appear to lend itself well to measurement and analysis. It is therefore desirable to develop a surrogate for the dictionary definition of risk that is amenable to quantification. For it to be intuitively pleasing it must measure, either directly or indirectly 'the chance of injury, damage or loss, so that it may be used synonymously with the word risk (Francis 1986).

More generally in analysis, risk is defined as being described by a known probability distribution of a particular event occurring, in contrast to uncertainty where the probabilities are unknown, with the surrogate measure of risk involving the variability, or some monotonic transformation, of that distribution. The greater the variation of that distribution then the greater is the risk of that particular event not occurring (Van Horne 1981).

Although risk and uncertainty are frequently used interchangeably, no distinction is made between the two in this research. They are conceptually seen as describing the same probability of loss or harm. The elements of doubtfulness, fickleness or changeability that characterise uncertainty can be described just as much by a subjective probability distribution as can risk by an objective probability distribution (Francis 1986). As such, uncertainty is not recognised as being distinctly different from risk in the context of this thesis.

There needs to be a clear distinction between the risk of a source stimuli as an event occurring, and the risk of an adverse or harmful result occurring as a consequence of that source event. For instance, we can refer to 'the chance, probability or risk' of an event such as a drought or price crash occurring, or we can refer to 'the chance, probability or risk' that the occurrence of a drought or price crash will be harmful, where harm itself is considered the event.

Two aspects of the verbal definition for risk require clarification and definition. First 'chance' and second 'damage, loss or harm'. Chance is easily interchangeable with probability. In the context of the risk definition, chance, possibility, probability and odds are synonymous. Therefore the definition can be altered to 'the probability of damage, loss or harm', where probability is indeed either objectively or subjectively quantifiable.

In order to define 'harm' as an event, in the context of the thesis definition for risk, one must pull together those aspects of both agricultural and banking risk, discussed earlier, that are identifiably common to both.

It has been indicated earlier that the level of agricultural risk inherent in the occurrence of a source stimuli varied according to the impact that resulted. Regardless of the source of that risk, the final impact of any source component of agricultural business risk is ultimately reflected in either a change in farm revenues or a change in farm expenditures, or both. An adverse impact would obviously consist of either a decrease in farm revenue or an increase in farm expenditures. More precisely, an adverse impact would be reflected in a decrease in farm profit, the magnitude of which essentially determines the magnitude of the financial risk inherent in that specific agricultural activity, through the effect that the decrease in profits has on farmer equity. Therefore the ultimate impact of any combination of business risk stimuli is itself a source of financial risk.

However a small decrease in farm profits, or equity, is not harmful if 'sufficient' profit and equity remain after the impact, but is harmful if little equity existed beforehand, and a financial loss resulted rather than a profit, thus causing negative equity. In this situation the financial loss would not only increase the default risk the bank first undertook in financing the operation, in terms of interest default on both long and short term loans and repayment of seasonal loans, but would also threaten the security underlying the total financing of the farm. If the situation were such that default did in fact occur then the situation would also constitute harm to the bank.

The commonality between agricultural and banking risk is therefore identified within one financial risk component of both entities, i.e., the profit component of farmer equity within the financial risk inherent in the farm operation and the default risk component of the financial risks inherent in banking. Joint, or common 'harm' can therefore be defined as 'zero or negative farm profit'.

The thesis definition of risk, which is also tantamount to a definition for default risk, within the context of short term seasonal financing of agricultural production, then becomes 'the probability of zero or negative farm profit', where zero or negative farm profit is assumed to be a jointly harmful event.

One more component of the thesis definition argument needs to be examined before a true link between farm financial risk and bank default risk can be established. It is contained within the definition of 'farm profit', and the implications associated with the accounting definition of farm profit.

1.3.2 FARM PROFIT AND THE CURRENT ACCOUNT

By far the largest proportion of trading bank gross operating revenue consists of interest income. The Bank of New Zealand group reported in its 1990 consolidated profit and loss statement that it received eighty percent of its gross revenue through interest earnings, and fifty percent of its total operating income from net interest income, net of interest payments (Bank of New Zealand 1990). Bank profit is therefore determined by the relative interest magnitudes of both assets, in the form of advances, investments and securities, and liability deposits.

Two characteristics of interest revenue are important in the context of risk. First, interest is essentially a cash revenue sourced as cash payments made by the lender from revenues derived by the activities the bank is financing. Second, the duration of that interest revenue is related to both the solvency of the lender, or his ability to continue servicing the loan, and the security underlying the loan in relation to its term.

From a bank revenue point of view, solvency would seem to be a more important component of minimising total default risk than would security. If the activity being financed is strong enough to provide, or guarantee, the ongoing servicing requirements of the loan, but has poor financial security in terms of that loan, then so long as that security improves over time as a consequence of the activity's strength, it would seem rational for the bank to continue its involvement.

Assuming that lender solvency is a dominant characteristic of continued bank involvement, and security is a dominant characteristic, or pre-condition, for initial loan provision, and both solvency and security are related via the financial strength of the activity, then it follows that from both the bank and farmer point of view, cash flow, or cash solvency, is the dominant criteria upon which default risk should be evaluated.

In terms of monitoring and measuring the default risk of bank involvement, then monitoring the current account is the only effective way of gaining information regarding the cash strength or solvency of the borrower. Current account is defined as the sum of all accounts, bank or otherwise, through which all cash transactions are made. For this reason, and because solvency is related to cash, the normal reporting format of farm accounting needs to be adapted to accommodate the thesis definition of risk.

The recommended format for farm accounting, as outlined by the New Zealand Society of Accountants 1985, is diagrammatically abbreviated in Figure 1.1.

With the emphasis on cash flow, the need for re-defining 'farm profit' becomes apparent when one notices the combination of

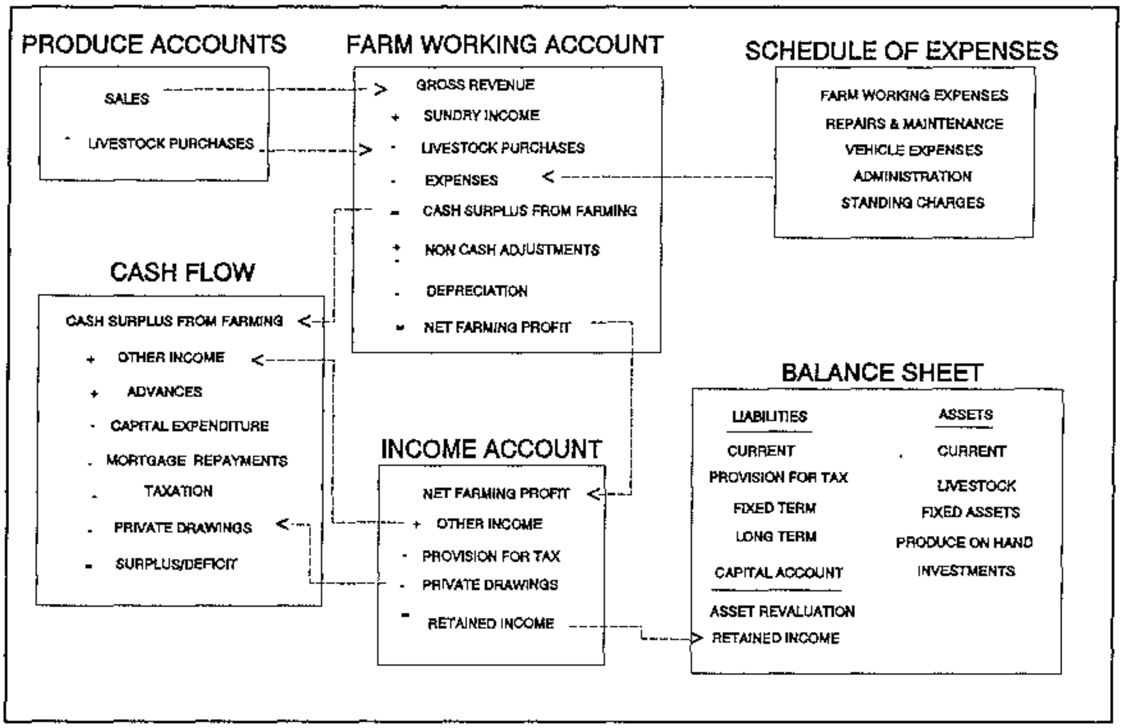


Figure 1.1 Farm Accounts

tangible objective cash items with comparatively intangible subjective values throughout these accounts. Cash surplus from farming is a true cash definition which is adjusted according to changes in livestock values and depreciation to derive net farming profit.

The cash flow statement, although a true cash definition, incorporates injections of loan capital or advances, which confuses the issue of cash solvency in terms of interest payments derived from the 'strength of the enterprise'. Cash injections of borrowed capital, if used to repay seasonal debt, also confuses the issue of seasonal finance secured by future farm production. The use of cash injections, whilst constituting a cash flow transaction through a current account, also increases both the liabilities and assets of the balance sheet (depending on whether or not capital is purchased with the injection) therefore directly influencing the equity security issue within default risk.

By separating out all intangible subjective components of the farm accounts, as well as removing external cash injections other than earned revenue, we are left with a net cash position that is related more to the capacity of the individual and his farm to earn sufficient revenue such that he can be defined as being 'productively solvent and secure'. As such we are attempting to separate out those cash components of the farm operation that directly relate to 'farm solvency' and distinguish between lending to achieve solvency and earning to achieve solvency.

The relationship between solvency and security is identified diagrammatically in Figure 1.2. Assuming that security is defined as percentage equity, where equity represents the proportion of the capital account to total assets, then one can clearly see from the diagram how important a positive cash flow is in relation to equity and security.

The diagram is not meant to show the relative magnitudes of the effect of changes in any of the accounting components, nor does it show the off-sets with regard to changes in 'below the dashed line' intangible components caused by changes in the 'above the dashed line' tangible components. It merely tries to establish the relationship between tangible cash components

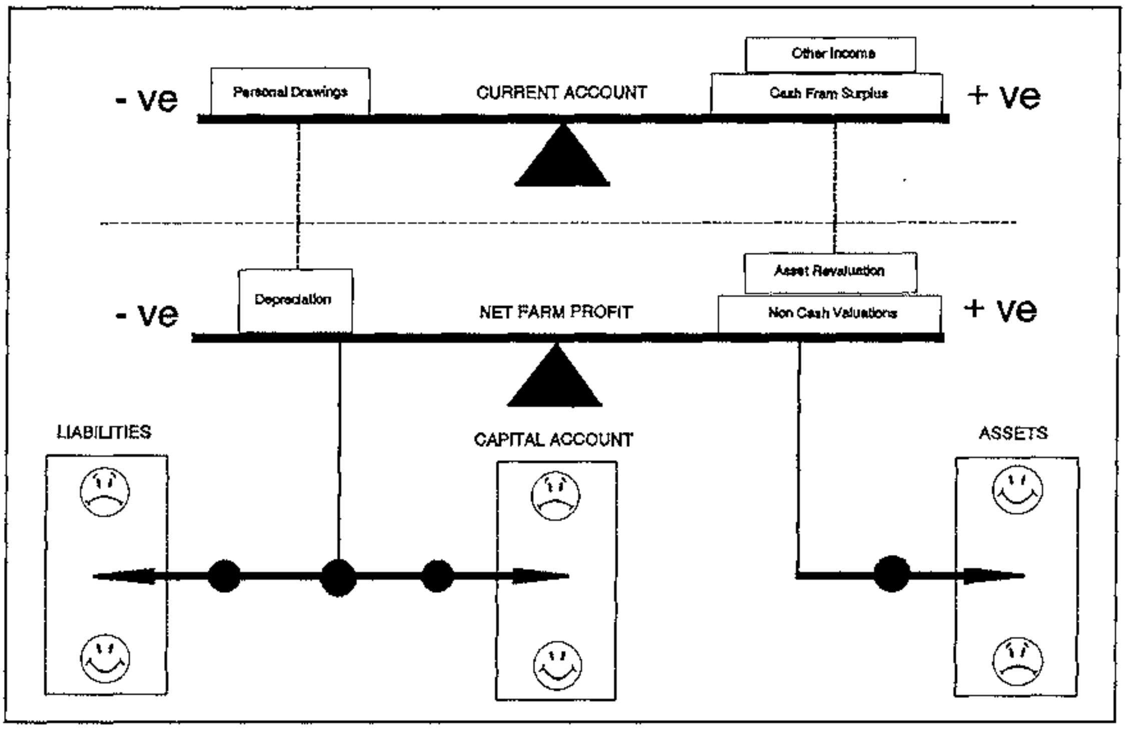


Figure 1.2 Cash balance

and equity. The diagram also indicates the sensitive balance between asset values, and adjustments, farm cash flow and cash profit, and the balance sheet components.

In the context of the thesis definition for risk, we define risk inherent in the provision of seasonal finance to agriculture as being ' the probability of incurring zero or negative net cash returns' where net cash returns are defined as being 'the sum of all revenues earned by the farmer on the farm less all cash expenses and payments made by the farmer'. The risk so described refers to the risk of insolvency and the inability of a farmer to service the sum total of all borrowing. This definition links together the two common financial risk components of agriculture and banking.

1.4 SEASONAL CREDIT AND THE NEED FOR RISK QUANTIFICATION

1.4.1 THE EQUITY PROBLEM

Amongst the many repercussions on the rural sector of the removal of agricultural support policies during the early 1980s, was a general 'across the board' loss of farmer equity. With government support and protection having been capitalised into land values prior to 1984, their removal almost immediately decreased land values and, as a consequence, farmer equity levels (Pryde 1987). Compounded by decreased product prices, many farmers found themselves sustaining and servicing debt levels with little or no underlying security.

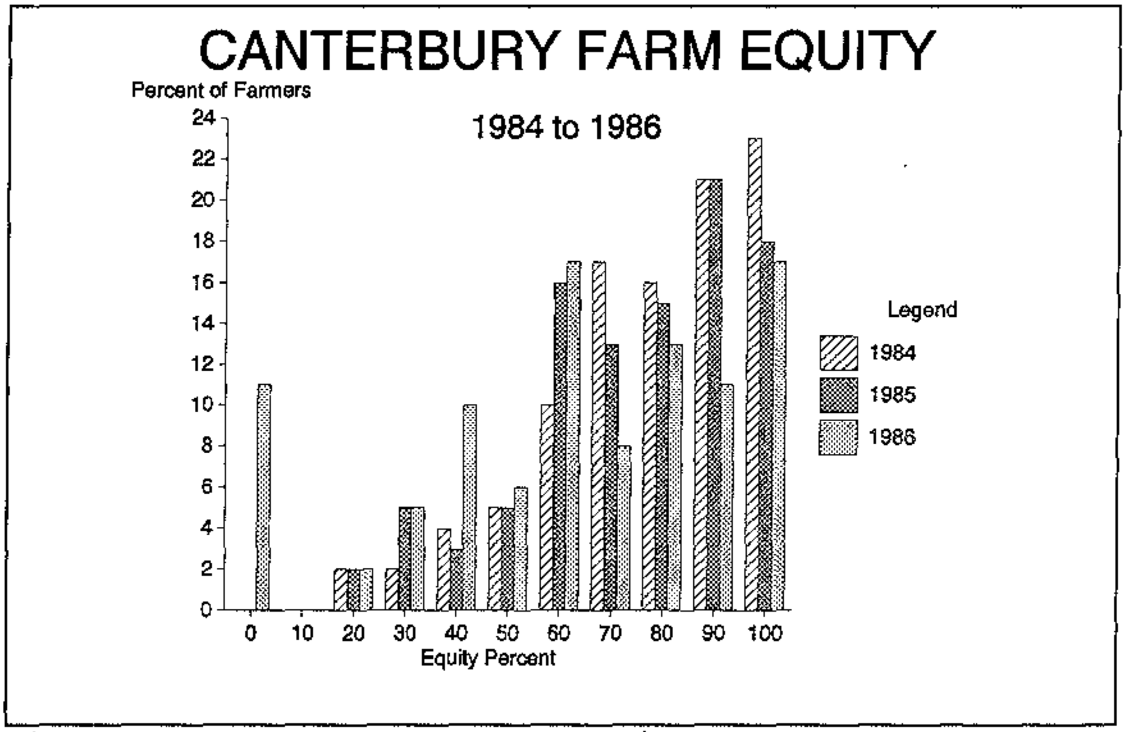


Figure 1.3 Canterbury farmer equity Source: Pryde 1987.

Evidence supporting a general loss of equity can be found in Pryde's 1987 analysis of the equity levels for a sample of Canterbury farmers. Figure 1.3 shows that during the period 1984 to 1986 inclusive, the sample distributions of farm equity changed from left skewed exponential to bimodal. The interesting feature of his analysis is that although all farm equity levels had dropped, by 1986 the distribution showed that no farmer in the survey had an equity of between 10% and 20%, and 11% of farmers had equity levels of less than 0%. Pryde's analysis also shows that in the period 1983/84 to 1985/86 the percentage of Canterbury farmers with 50% or less equity had grown from 13% to 35%.

Pryde's analysis is also supported by the 1984/85 to 1988/89 equity distributions, displayed in Figure 1.4, of a national sample of sheep and beef farmers taken annually by the New Zealand Meat and Wool Board's Economic service.

This distribution illustration also indicates a sudden 1986 increase in the percentage of sheep and beef farmers with zero or negative equity levels. For the 1984/85 season only 0.2% of sheep and beef farmers were in this category. By the end of the 1985/86 season, the percentage of farmers in the negative equity category had increased to 3.7% of the sample, inferring that some 800 sheep and beef farmers, among approximately 22,000 sheep and beef farmers at that time, had absolutely no financial equity security, and the percentage of sheep and beef farmers with 50% or less equity had grown from 7.9% to 20.8%.

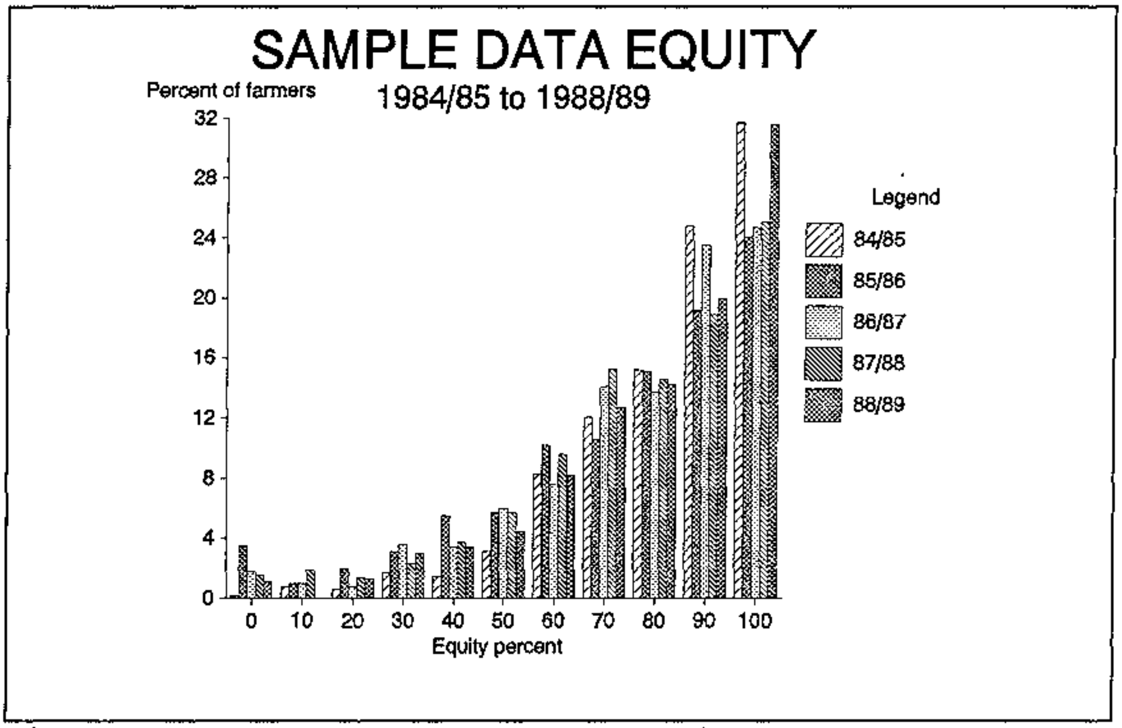


Figure 1.4 Sheep and beef Farmer Equity Source: M.W.B.E.S

By the end of the 1988/89 season the percentage of negative equity farmers had improved to 2%.

In discussions with numerous financial advisors, Pryde found that in assessing the financial situation of farmers, they now paid no attention to averages, and less attention to equity. Farmers are categorised into 'boxes' that depict distinct situations (Pryde 1987). In a typical cross-section of Canterbury farmers, 6% were in deep financial trouble, and had been for at least ten years, 21% were in trouble but probably could recover under favourable conditions, 49% were described as struggling, having been caught by their level of borrowing, and 24% were described as 'very sound'.

If we consider just the first 'box', 6% in serious trouble, in just Canterbury alone, there were between 300 - 400 farmers in this predicament (Pryde 1987). Over the country this may conservatively be estimated at between 4000 - 5000 farmers. Further, the total agricultural debt at 1987 was estimated to be about $8 billion. With an average indebtedness of approximately $150,000, this equates to a total of between $600 million and $750 million of at risk farm debt, and this is for only 6% of the total farmer population. It is unlikely that by 1991 the situation is much improved, despite the improvement noted in Figure 1.4. Note that the distributions also infer that banks must indeed be currently financing at least 1.5% of farmers with no equity security.

Although Pryde's research describes only the Canterbury experience, it seems probable, according to figure 1.4, that similar equity situations occurred throughout New Zealand. Pryde's 1987 survey of financial advisors and institutions revealed a profile of the type of farmer affected by financial difficulties. His list of characteristics, abbreviated below, would appear to be generally applicable.

Characteristics of farmers in financial difficulty:

1. The standard of resource management is generally poor.
2. Working expenses exceed a certain proportion of income.
3. A lack of economies of size and scale.
4. Excessive levels of personal drawings.

1.5 CHAPTER SUMMARY

The current environment within which both farmers and bankers conduct their business demands a more direct responsibility for the management of risk. Given the unique business risks inherent in agricultural activity, plus a need for bankers to avoid involvement in farm plans that increase the risk of loan servicing and repayment default and the adverse consequences of default occurring, specific agricultural risk needs to be measured before risk management strategies can be developed and implemented.

The specific agricultural risk to be measured is defined as 'the probability of a farmer client incurring zero or negative net cash returns'. A client that conducts a farm activity which results in a negative cash position after one season of operation, is deemed to be unable to totally service existing debt. A pre-condition of servicing is that funds available for servicing must be sourced from earned activity and not borrowed.

Net cash returns are defined as consisting of all earned gross cash revenues less all farm and personal cash expenditures. The proposed cash transactions, or farm budget, conducted through the farmer's current account, excluding injections of borrowed capital, will be utilised in the quantification of default risk.

Measuring the probable inability for an activity to service debt constitutes a warning mechanism with regard to lending secured by farmer equity. Avoidance of the repercussions, on both banker and farmer, of loan interest and repayment default underlie the objectives of such a measurement.

1.6 THESIS OBJECTIVES

The first objective of this thesis is to quantify the default risk of seasonally financing agricultural production according to the definition 'the probability of zero or negative net cash returns'. It is intended that the probability be directly estimated utilising the probability modelling technique of logit analysis, and applied to the dominant components of sheep and beef farming.

The second objective is to concurrently develop a method for combining multiple farming activities such that the resulting risk measure refers to an individual farmer client involved in any combination of those activities.

1.7 THESIS ORGANISATION

Chapter One has described the background and associated need for the measurement of a specific type of risk common to both farmers and bankers, and has identified and defined that risk.

Chapter Two evaluates traditional risk measurement techniques for their general suitability as a risk measure. Chapter Three presents and describes probability modelling techniques.

Chapter Four describes and defines the variables used in the modelling process, as well as describes the data used in the construction of the variables. Chapters Five through Seven specify the logit model and report the results of model building and testing.

The thesis concludes with Chapter Eight, which discusses the results, strengths and weaknesses of the model, as well as a framework within which the model could be utilised. Limitations and problems encountered in this analysis will indicate areas for further research.

Chapter Two

RISK QUANTIFICATION METHODS

2.1 TRADITIONAL METHODS OF RISK MEASUREMENT

It has long been accepted that measurement of the risk involved within any investment consists of the objective or subjective variability about expected investment outcomes. First and second moment analysis about an expectation, including variations such as standard deviation, semi-variance and covariance, form the basis of the majority of recognised risk evaluation techniques and models. Since the 1952 Markowitz publication 'Portfolio Selection', and the birth of portfolio theory, researchers have generally concentrated on methods of selecting investment or activity options with risk, having first of all accepted variability as a surrogate measurement of risk.

With the objective of minimising risk concurrently with the maximisation of returns, methods of activity selection logically centred on the tangency between either a utility or objective function, and what is commonly known as an E-V frontier or opportunity curve. The E-V frontier plots the various returns (outputs) of all options on the Y axis against the variance about those returns on the X axis. Tangency with a preference/utility/objective function indicates the optimal trade-off between returns and risk, and identifies the most 'suitable' option of those plotted.

Interest evolved into the structure of the investor's utility curve and the nature of selection behaviour under risk. Much literature on risk analysis concentrates on either the behavioral aspects of decision makers to identify their level of risk aversion, or assumes a level of risk aversion or utility curve and identifies the optimal activity selection under either a linear or quadratic programming framework (Bardsley, Borch, Lambert, Scott, Robison, Taylor, Pope). All have in common, with some exceptions, variability or covariability as their quantification of risk.

Those that recognised the limitations of variability and the E-V frontier utilised a Bayesian or Bernoullian probability framework for their analysis (Rae 1971). The associated development of decision trees and the selection of activities based on an individual's appraisal of outcomes can ignore an objective risk measure by relying on the individual to utilise a 'black box utility function' in selecting his/her activity on the basis of maximising the subjective probability of an expectation actually occurring. The subjective probabilities utilised in such an analysis are themselves surrogate risk measures.

Various mixtures of sensitivity or variability of subjective or objective probability distributions as risk measures are found in models such as CAPM, MOTAD, APT, and MONTE CARLO, which are outlined in this chapter. Very few models utilise the third and fourth moments of a distribution. Most make the assumption of normally distributed returns and normally distributed risk variables. This chapter will examine the various methods of risk quantification for their suitability to agricultural risk evaluation.

2.2 INCOME VARIATION, SENSITIVITY AND THE E-V FRAMEWORK

2.2.1 VARIANCE

Variance, and monotonic transformations such as standard deviation, standard error, absolute deviation, coefficient of variation etc, are generally utilised as surrogates for measuring risk. Generally used within a portfolio selection framework, variance about an expected outcome attempts to convey the strength of uncertainty one should feel toward the expectation. The greater the variance, or dispersion, the more unlikely is the expected outcome to eventuate.

Given a farm firm that produces i products, then the farms expected net returns will be:

$$E(r) = \sum_{n=1}^{i} q_n x_n = q_1 x_1 + q_2 x_2 + \ldots\ldots q_i x_i \qquad (2.1)$$

where q represents net returns per unit of x produced. Alternatively, expected net returns may be the sum of the products of the various one-period net returns times their probabilities, where x_n denotes the nth net return from the probability distribution, q_n is the probability that the nth net return occurs, and there are i possible net returns (Francis, 1986).

Variance is then calculated according to whether (a) the expected returns represent the sum of individual activity returns within one-period, or (b) the expected return is the various total farm returns times their probabilities of occurring.

In the case of (a), we calculate the total variance function assuming that the q_n's are random variables with means, q_n, n=1...i, variances and covariances σ_{nm}, n=1....i, m=1...i (when n=m, $\sigma_{nm} = \sigma_n^2$) i.e. the net returns from each activity have an expected value or mean, variance about the mean, and covariance with the returns from other activities within the total expected returns equation.

The variance of the total farm net returns V(r) can be expressed as a quadratic function of the x_n's and the variances and covariances of the q_ns (Stovall 1966):

(2.2) $$V(r) = \sum_{n=1}^{i} \sum_{m=1}^{i} \sigma_{nm} \, X_n X_m$$

The first and second partial derivative of (r) with respect to x_n can be used to determine the marginal contribution of the nth activity to total net return variance. That is, a variance risk measure can be assigned to each individual farm production activity.

(2.3)

$$\frac{\partial V(r)}{\partial X_i} = +2 [\sigma_{1n} \, X_1 + \sigma_{2n} \, X_2 + \ldots\ldots\ldots \sigma_{(n-1)n} \, X_{n-1} + \sigma_n^2 \, X_n]$$

(2.4) $$\frac{\partial_2 V(r)}{\partial X_n^2} = 2\sigma_n^2$$

Since the second derivative is always positive, the sign of the first derivative determines how incremental changes in the level of the nth activity affect total return variance. With the addition of an activity, or an increase in any existing activity, the sign of the first derivative shows whether the addition of a unit of the new activity will result in a higher or lower income variance. With $X_n = 0$, the last term in the first derivative vanishes, reducing to;

(2.5) $$2 \sum_{n=1}^{i-1} \sigma_{ni}\, x_n$$

i.e., twice the sum of the covariance of the income from the additional activity with the incomes from each existing activity weighted by the respective activity levels. If the above equation is negative, incremental increases in X_n will reduce total income variance.

In the case of (b), variance is calculated directly from the probability distribution of likely net returns. ie:

(2.6)

$$\sigma^2 = \sum_{n-1}^{i} q_n\, [x_n - E(r)]^2$$

$$= q_1\, [x_1 - E(r)]^2 + q_2\, [x_2 - E(r)]^2 + \ldots\, q_i\, [x_i - E(r)]^2$$

The resulting variance of a farm firm's expected returns, calculated according to the probabilities of different return levels, attempts to measure the risks inherent in such distributions, subjective or otherwise (probabilities utilised may be based on actual past returns).

2.2.2 SEMI-VARIANCE

Variations on the variance theme include conversion to standard deviations, standard errors, mean absolute deviations and semi-variance or other monotonic transformations of variance. Use of a semi-variance measure at least acknowledges the fact that distributions of net returns are unlikely to be normal.

If risk is defined as the probability of loss or harm, it seems more logical to measure risk by the area in the probability distribution that is below its expected return. Tsiang indicates that if the third and fourth moments of a distribution are not significant, then second moment analysis is sufficient. This merely states the obvious in that if a distribution is normal, i.e., insignificant skew and kurtosis, then variance and use of the empirical law are sufficient to make statements and give an indication of risk. It is not important whether variability of returns (risk) is measured on one or both sides of the expected return (Francis 1986).

Semi-variance, which measures variance below the expectation, is calculated

$$SV(r) = \sum_{n-1}^{i} q_n \, [\, k_n - E(r) \,]^2 \tag{2.7}$$

Where k_n is the nth below average return of the distribution.

The question therefore seems to be: are net returns, or the variables that construct net returns, normally distributed or skewed? If returns are skewed, then variance or any monotonic transformation cannot accurately reflect risk.

2.2.3 SINGLE-INDEX COEFFICIENT VARIANCE

A useful variation on the variance theme can be found in the single-index model of risk (Sharpe 1963,1970) adapted for use within a quadratic programming framework by Collins and Barry (1986). Their objective was to develop a single-index risk measure, based on single-index parameters and computational simple methods for farm risk planning.

The single-index model provides for the measurement of individual activities within a multiproduct farm. But again it contains no more than a variance-covariance matrix. The model is based on the assumption that each activity's return (r_i) is linearly related to some common factor (r_m) and to random elements e_i.

(2.8)

$$\overline{r}_i = \alpha_i + \beta_i \overline{r}_m + \overline{e}_i$$

The r_m variable should be any factor thought to be the most important single influence on returns (Sharpe 1970). In finance, GNP or stock market indices can be used. Within agriculture, climate indices or land area might be appropriate. The beta coefficient represents the single-index measure of risk. Collins assumes a region of homogenous land has N crop production activities with expected returns r_i (i=1,...N) to risk, management, and capital. Further, he assumes that variable r_m is a generalised measure of the region's income. Its values would depend upon the region growing conditions and prices for resources and products.

Accordingly, beta coefficients measure the systematic volatility of the respective crop returns. A crop with beta = 1 would on average experience the same systematic volatility as the average of all crops in the region, and would 'follow the market'. A crop with beta = 2 would have double the systematic volatility of the average of all crops, and so on. Using Ordinary Least Squares the variance of r_i is

$$(2.9) \qquad Var(r_i) = E\ [\alpha_i + \beta_i r_m + e_i - E(\alpha_i + \beta_i r_m + e_i)]$$

$$= \beta_i^2 \sigma_m^2 + \sigma_{ei}^2$$

The variance of an activity therefore consists of two parts. One part ($\beta_i \sigma_m^2$) measures the variability common to all activities, and the second part (σ_{ei}^2) measures an activity's unique variability, i.e., the nondiversifiable (systematic) and diversifiable (nonsystematic) risk respectively.

The farm business that produces M crops (M <= N) where X_i is the proportion of land in the ith crop will have expected returns of

(2.10)
$$r_p = \sum_{i=1}^{M} X_i r_i \quad where \quad \sum_{i=1}^{M} X_i = 1$$

and variance

(2.11)
$$\sigma_p^2 = [\sum_{i=1}^{M} X_i \beta_i]^2 \sigma_m^2 + \sum_{i=1}^{M} X_i^2 \sigma_{ei}^2$$

where the variance of a single crop activity, and thus its risk measure is

(2.12)
$$\sigma_i^2 = (\beta_i \sigma_m)^2 + \sigma_{ei}^2$$

which reduces to

(2.13)
$$\sigma_i^2 = (\beta_i \sigma_m)^2 \qquad \sigma_i = \beta_i \sigma_m$$

if the activity is produced by a well diversified farm.

Thus the set of single-index beta coefficients approximates the variance-covariance matrix and serves as a measure of risk in agricultural portfolio analysis. Collins believes that the choice of r_m is not critical.

2.2.4 VARIATION, THE BETA COEFFICIENT AND CAPM

In agriculture, the traditional portfolio choice problem is based on the total variance of the farm plan relative to expected returns. The contribution that each farm activity makes to the total variance of the farm plan has received relatively little attention (Turvey, Driver 1987). As has been outlined in the single-index beta model (Sharpe), the farm sector portfolio reflects only that risk that is common to all activities, i.e., nondiversifiable risk. For individual activities, there are also two types of risk - nondiversifiable because of the correlation with the total farm plan, and the diversifiable, or non-correlated element of the activity.

Generally, a large diversified portfolio of investments is considered to be free of any diversifiable risk, i.e., all the risk is systematic. But within the individual investments i, each have the two component risk elements. Therefore the degree to which i's risk is systematic relates to the degree of correlation that exists between i and the total portfolio over the same time horizon. The relationship is captured by what is known as the characteristic line;

(2.14)

$$R_{it} = \alpha_i + \beta_i R_{mt} + e_{it}$$

where β_i is the beta coefficient for investment i, R_{it} is the return on i, R_{mt} is the total return of the portfolio, over time horizon t.

ß is then the relative measure of i's systematic risk. Since it is a measure of the correlation between R_{it} and R_{mt}, it is the predicted response of R_i to changes in R_m.

(2.15)
$$\frac{\partial R_{it}}{\partial R_{mt}} = \beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)} = \frac{r_{t,m}\sigma_i\sigma_m}{\sigma_m^2}$$

where $r_{i,m}$ = the correlation between R_{it} and R_{mt}. The systematic component of i is defined as $r_{i,m}\sigma_i$ and the nonsystematic component $(1 - r_{m,t})\sigma_i$ (Levy, Sarnat 1982)(Turvey). When the market is in equilibrium, the expected return on i is directly related to the systematic risk of i. This relationship is captured by what is known as the market security line

(2.16)

$$E[R_i] = R_t + \beta_i (R_m - R_t)$$

where R_t is the return on a risk free asset, R_m the mean return of the portfolio, $(R_m - R_t)$ the risk premium, where $\beta_i(R_m - R_t)$ establishes the premium above R_t required to hold security i, which has as its beta risk measure β_i (Turvey, Driver 1987).

Using the equations for both systematic and nonsystematic risk, each component can then be calculated on a cash basis, and choice of activity made according to deviation away from the market security line. Figure 2.1 represents a hypothetical market security line for a range of agricultural activities.

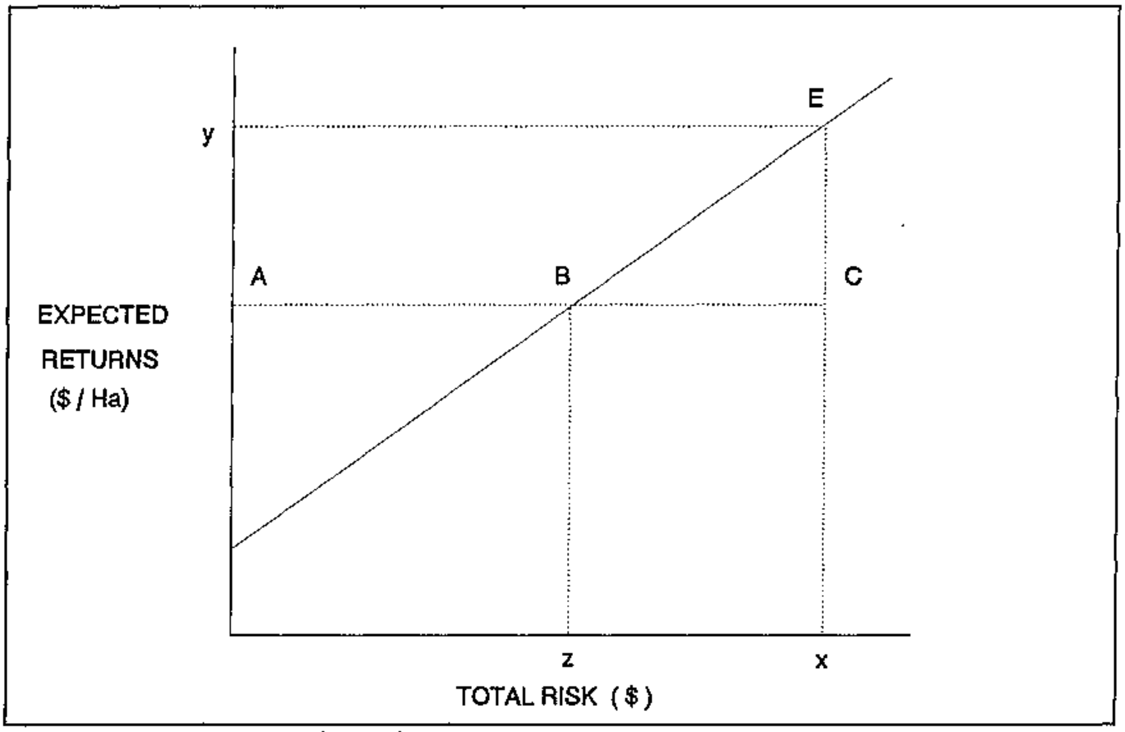


FIG 2.1 Portfolio Risk

Assuming that point E represents one activity with a total risk in dollar terms of X, all of which is assumed to be systematic, then one would expect to receive $Y per hectare. However, using the equations for systematic and nonsystematic risk above, one finds that systematic risk is only \$Z, ie. AB is systematic risk, while \$ZX (BC) is nonsystematic or diversifiable. Point B then represents an expected dollar return \$A, for accepting \$Z per hectare worth of systematic risk. Therefore the actual \$ risk measure for activity i equals \$BC.

This method identifies a beta risk measure for activity i as well as identifying a dollar risk value. Turvey and Driver conclude that using this method, Beta and a CAPM framework can be used to identify the specific activity systematic and nonsystematic risk. Further they content that this method also identifies whether or not farmers are actually compensated for the level of systematic risk they accept.

Those within investment management have found that the beta coefficients generated have tended to be poor, i.e., low R^2 or poor fits. Brealey (1984) and Gooding (1978) believe investment risk to be a multidimensional concept. They found also that although returns correlated significantly with beta, they did not correlate with total portfolio risk. Further, the perceived beta measure and perceived financial leverage were also uncorrelated.

An application of CAPM to New Zealand agriculture was attempted by Narayan and Martin (1990). They concluded that generally the choice of a market portfolio is crucial to the outcome of the model, and although useful to decision-makers, is therefore extremely sensitive and should be used with caution (Narayan, Martin 1990).

2.2.5 ARBITRAGE PRICING THEORY

APT or arbitrage pricing theory (Ross 1976), is a model that differs from CAPM in that it skips the step of how investors can construct efficient portfolios, and assumes that each stock's return depends on several independent influences or factors.

(2.17)

$$r_{it} = \alpha_i + \beta_{i1}F_{1t} + \beta_{i2}F_{2t} + \ldots\ldots\ldots + \beta_{ik}F_{kt} + e_{it}$$

is the general form of what is a single asset, time-series return-generating model, where r_{it} is the one-period return from the ith asset in time period t, α_i the expected return for an asset if all risk factors have a value of zero, F_{jt} the jth risk factor that impacts upon the assets return, j = 1,2 ...k different risk factors with expectation of zero, ie $E(F_{jt}) = 0$, β_{ij} a sensitivity indicator or factor loading that measures the responsiveness of the asset to F_{jt}, and e_{it} the residual return unexplained, [$E(e_{it}) = 0$]. All factors must be independent (orthogonal) with respect to all assets, i.e., $E[e_i\{F_j - E(F_j)\}]=0$ for all i and j.

The betas in this example are again proxy risk measures similar to the single-index model and CAPM. The APT model however goes one step further; it establishes a risk premium that would induce investors to assume a one-unit increase in beta risk. i.e.

(2.18)

$$E(r_i) = \lambda_0 + \lambda_1\beta_{i1} + \lambda_2\beta_{i2}\ldots\ldots +\lambda_k\beta_{ik}$$

The λ_j coefficients thus measure the market price of risk for whatever risk is measured by β_{ij}. λ_0 represents the return on a riskless asset, if such an asset exists. Subtracting the intercept term from $E(r_i)$ indicates the risk premium form for that portfolio.

If either beta in a two factor model equals zero, and the other equals 1, then the APT equations reduce to the CAPM equation.

$$E(r_i) - \lambda_0 = \lambda_1\beta_{i1}$$

(2.19)

$$E(r_i) = \lambda_0 + \lambda_1\beta_{i1}$$

$$with\ \lambda_1 = E(r_m) - \lambda_0$$

$$with\ \lambda_0 = R$$

This interpretation shows that CAPM is merely a special case of

APT. Or, APT is a logical extension of CAPM. In either case the risk measure is beta, but risk can be conveyed conceptually better, within APT, by expressing it as a dollar premium. Investors might possibly relate better to a dollar 'risk price' than simply a beta risk measure.

2.2.6 BETA AS A MEASURE OF RISK

Relatively little dispute appears in the literature regarding the suitability of beta coefficients as measures of risk, although their accuracy and usefulness would appear to depend on the exogenous variable or variables used in their approximation. APT at least identifies a need to use those factors directly contributing to the risks of any stock or activity within a portfolio. Still, the methods employed by CAPM and APT rely on variance as a proxy measure, by deriving a coefficient that is in effect a measure of the variance and correlation of the stocks variability in relation to either the 'market' or other chosen factors.

Estimates of correlation coefficients are obtained from betas by noting that

(2.20) $$Cov(ij) = \rho_{ij}\sigma_i\sigma_j = \beta_i\beta_j\sigma_m^2$$

All of the beta methods assume that the only correlation that exists between securities arises because of common correlation with the market. This is a simplified assumption in that it ignores additional sources of correlation present, such as that which might arise from specific industry conditions. It has been shown that ignoring these industry specific correlations leads to a downward bias in the correlation estimates (Elton, Gruber Urich 1978). Further, they conclude in their study that overall mean is a preferred method of forecasting correlation coefficients, and therefore betas, in comparison to the best of the time series beta techniques. Their results show that a naive average of correlation dominates all other techniques at a statistically significant level for all periods sampled. Their conclusions basically invalidate betas as surrogate risk measures, due to the large errors in their estimation and the simplicity of their underlying assumptions concerning covariance and correlation with the market.

It seems that any exogenous market index variable on its own is not sufficient to explain the variation of individual securities. APT attempts to correct for this by including several variables, but again their selection may itself be a major task given the requirement that all factors used must be independent, if total variation of a single security is to be explained.

2.2.7 ABSOLUTE DEVIATION - MOTAD

A transformation of the variance theme is found in what is generally known as the MOTAD model. MOTAD minimises total absolute deviation rather than variance, or minimises the sum of the absolute values of the negative deviations using a linear programming algorithm (Hazell, 1971). The objective of the model is to develop risk efficient farm plans, much in the same way that EV/utility analysis attempts to do, by parameterising an income constraint from zero to its maximum value, thus tracing out the EV frontier.

A typical formulation of the MOTAD model may look like

Maximise : $L\overline{d}$

Subject to : $A\,X \leq B$

$DX + I\overline{d} \geq 0$

$C'X = \lambda$

(2.22) $X,\ \overline{d},\ \lambda \geq 0$

where X, A, B, and C represent activity labels, resource requirements, resource availabilities, and gross margin expectations, respectively (Mapp, Hardin, Walker, Persaud 1979). The gross margin expectation is the mean of the series.

Element D represents a matrix of deviations between the observed gross margin and the gross margin expectation. The vector d represents yearly total negative deviations summed over all risky activities. Ld represents the summed total negative deviations over all years. λ is a scalar used to represent the income constraint. The efficiency frontier is traced out by parameterising this scalar, and the solution set indicates the optimal activity levels under different risk scenarios, specified as input data for the model, and tests the feasibility of certain farm plans, under risk conditions.

The model does not implicitly derive risk measures as such, but instead assumes risk is described by absolute deviation between observed and expected gross margins, and minimises the deviation.

2.2.8 SIMULATION AND SENSITIVITY - MONTE CARLO

Simulation and sensitivity analysis is generally useful for 'what if' questions relating to the impact upon cash flows and net returns of changes in various variables responsible for producing the net return. The Monte Carlo method of analysis utilises both probability distributions and sensitivity analysis, to construct a final net returns distribution from which variance and deviation can be estimated. The method overcomes the complicated procedure of mathematically combining independent distributions to derive the four moments of the resulting distribution.

For example, consider a farm production unit that produces n cash crop products i = 1....n. Its net cashflow return function may look like

$$E(r) = \sum_{i=1}^{n} [A_i(Y_iP_i - V_i)] - TC \tag{2.23}$$

where A_i represents the area of crop i, Y_i the per hectare yield of crop i, P_i the per unit (yield) price received for crop i, V_i the per hectare variable costs of crop i, ((Y_iP_i - V_i) would represent the per hectare gross margin of crop i) and TC representing the total fixed costs of the unit.

Each variable is independent, a constraining requirement of Monte Carlo analysis, but Y_i and V_i are correlated. In addition, each variable has a probability density function, either objectively derived from historical information, or subjectively estimated using a triangular distribution or some other method. Further, the distributions for A_i, P_i, V_i, TC and Y_i are finite, i.e., $0 > A_i <= 100$ (the total area of the farm), and $0 > Y_i <= x_i$ (x_i representing the maximum possible yield for crop i). Prices and costs can neither be less than zero nor greater than some predetermined level.

Monte Carlo then randomly samples point estimates from each variable distribution and completes the net return function using the selected estimates. After completing a predetermined number of iterations, the method has constructed a random distribution of possible returns. One then has the choice of using either the mode or the mean of the resulting distribution as an expectation of net returns, and appropriate statistics calculated from the distribution. However, other than a potentially more accurate distribution of returns, the decision maker is still confronted with variance and the like as measures of risk even if the variance in this case is likely to be a better reflection of the combination of all underlying variances (Brealey,Myers 1984).

Although the procedure would appear to be straightforward, problems arise in identification of the appropriate distribution functions from which samples are drawn, and specification problems arise in terms of identifying the appropriate final distribution. The Monte Carlo method can only accept moment measures of the underlying distributions. In combining these various distributions, parameter estimation problems and stochastic dominance problems arise, in attempting to measure the moments of the final distribution. Distribution forms and appropriate distribution functions need to be identified before estimation procedures can begin. Pope and Ziemer (1984) considered normal, log-normal, and gamma distributions in their search for stochastic efficiency using various estimation techniques, i.e., the mean-variance rule, maximum likelihood method, the empirical distribution function and appropriate ML methods for nonnormal cases. They conclude that for most risk efficiency analysis in agriculture, the empirical distribution function preforms better than other techniques, and that the popular mean-variance rule does not have general applicability since it is not robust toward nonnormality (Pope,Ziemer 1984).

2.2.9 VARIANCE AS A RISK MEASURE

Since the development of portfolio theory, based on the EV frontier and the separation theorem, discussion has centred not on the suitability of mean-variance as a surrogate risk measure, but as a method for identifying efficient choice, under risk conditions, relative to maximising utility.

Under a quadratic framework, the EV set derived is additionally constrained by activities needing to be divisible, nonnegative, and linear, i.e., outputs are linear combinations of the inputs (Robison, Brake 1979). Additionally, portfolio theory in most cases only considers price risk. Portfolio theory has been condemned as a consequence of the apparently restrictive assumptions. Samuelson (1970) and later Tsiang (1972) defended the use of portfolio theory and the mean-variance framework, once Tobin, one of the initial pioneers of EV analysis, admitted that mean-variance could only be used under special circumstances of known normal distributions and second order polynomial utility functions.

Tobin was forced to concede the weaknesses of EV through criticism from Borch and Feldstein, who contended that any system of upward sloping mean-standard deviation indifference curves can be shown to be inconsistent with the basic axiom of choice under uncertainty (Tsiang 1972). Borch believed that by the dominance axiom, any point on the curve could be preferred if its distribution and probability of gain was greater than any other, which contradicts the meaning of indifference curve. Feldstein used a log utility function to show that an ES curve for a risk averter need not be convex downwards, though upward sloping, suggesting that risk aversion might decrease as risk itself is increased beyond a certain extent.

Tsiang's response was to qualify the usefulness of EV by declaring that, with regard to utility, any utility curve that could be transformed into a polynomial through a Taylor expansion, and converged quickly enough, such that moments higher than two became insignificant, could be therefore represented by the quadratic utility function. He further asserts that mean-variance are suitable proxies for risk if the level of risk a decision maker faces is small, i.e., what is necessary for the E-V analysis to be a good approximation is merely that risk should remain small relative to the total wealth of the decision maker.

Further, this level of risk ought not exceed 10% of the individuals wealth. Indeed if risk is normally very small, then for a fair approximation for many problems, we may safely neglect terms higher than the second moment. Yet Alderfer and Bierman (1982) conclude in a study examining the importance of the third moment in the investor's decision process that the third moment, and possibly the fourth, are also relevant. Their experiments show that investors clearly use more information than just the first two moments. In addition, factors beyond the third and fourth moments were required to explain the choices made in their experiments. Probability of loss, certainty of payoff, and the maximum possible loss are also likely to influence choice.

Apart from the normality issue, Tsiang raises an important issue. To know how small the risk level is before we can safely use the E-V framework requires us to measure it. Does this mean that we can only use EV if the variances of activity options are small, and how small should they be? With regard to the normality issue, Tsiang himself admits that the assumption of normal distributions for all outcomes of risky investments or activities is patently not realistic; for it would rule out all asymmetry or skewness in the probability distributions of returns.

Two questions then: are agricultural returns normally distributed, and are the risks within agriculture relatively small enough to make E-V analysis a reasonable proxy for risk analysis and activity selection?

Richard Day (1965), in a statistical analysis of field crop yields concludes that normality and lognormality appear to be the exception rather than the rule. He noted the relationship between skewness, kurtosis and nitrogen application rates, suggesting extreme departures from normality across different crops. Further, he concludes that as a consequence, mode or median estimates of yield may be preferred to mean estimates both for forecasting and prescription purposes.

Steven Buccola (1987) showed positive correlation between skewness and kurtosis reduces the likelihood of associated decision errors from false imputation of normality. He cites the work done by Pope and Ziemer who show that use of EV methods in conjunction with nonnormal distributions leads to a relatively high rate of incorrect rankings among risky choices. By using Pearson's Skew and Kurtosis measures, Buccola shows the level of skewness and kurtosis for alfalfa and dryland wheat indicating nonnormality once series for prices, yields and costs had been whitened.

Thus the E-V framework as a decision method for agricultural production choice and bank lending policy, may be a risky method for agricultural finance application evaluation.

2.3 DISCRIMINANT FUNCTION ANALYSIS - CREDIT CLASSIFICATION

As was outlined in chapter one, risk to a banker can be categorised in terms of the risk of default due to a borrowers insolvency and bankruptcy, interest rate risk in terms of Government fiscal and financial policy, and market risk due to competition within the banking industry and the convertibility of deposit investments (Crosse, Cox). Given a farmer client, the bank is exposed primarily to default risk because of the unique risk nature of agriculture impacting directly on cash flows. Therefore, any agricultural risk directly measures the default risk faced by the bank, and from the bank's point of view, can and should be used as some form of credit measurement. Risk and credit are directly related.

The APT model gives a clue as to how a bank might measure risk and so derive a credit ranking. A multitude of factors influence the credit worthiness of any prospective client. These include not only the risks that underlie the specific activities the client is proposing, but also perhaps other physical or personal circumstances that might threaten the ability of the client to carry out those activities. Examples might be marital problems, disability or even death. In addition, personality characteristics such as extreme extrovert behaviour and a propensity to gamble, or conversely an introvert character with personality problems that may result in excessive drinking, can impact severely on the ability of the client to fulfil a financial bank contract.

It is not possible to identify every risk component and then measure each component within each prospective client, as well as the risk components of the activity the client is seeking to finance, such that a risk free credit decision can be made. However it is possible to identify some factors or characteristics that might act as indicators of the potential risks behind a credit decision. Alternatively, identifying and measuring only those objective risk components related to the proposed activity and then adjusting those measurements according to some subjective criteria in the formulation of a credit index might result in a more sound credit decision.

Discriminant function analysis attempts to do either of the above, in much the same way that the APT measures risk according to those factors thought to contribute to risk. The objective is to classify individual cases into predetermined categories according to independent factors thought to influence that classification (Peirson, Bird, Brown 1986). Assume that there are three categories of credit with which to assign any prospective client; good, doubtful and bad. Historical client data has classified known clients into the three categories according to loan performance. For each case within each category, data has been collated on three factors, X_1, X_2 and X_3, which might be actual net income, area of farm and rainfall respectively. On the basis of historical information, the bank wishes to predict membership into one of the three classifications based on the three predictor factors.

Because there are three classification groups, we need to estimate two discriminant functions; the first has the largest ratio of between-groups to within-groups sums of squares, the second, uncorrelated with the first, has the next largest ratio. In general if there are k groups, k - 1 functions can be calculated, in an attempt to explain as much variation as is possible. Generally the functions would take the form

$$D_{i1} = \alpha_{i0} + \beta_{i1}X_1 + \beta_{i2}X_2 + \beta_{i3}X_3 \tag{2.24}$$

$$D_{i2} = \alpha_{i0} + \beta_{i1}X_2 + \beta_{i2}X_2 + \beta_{i3}X_3$$

A case's values on both discriminant functions must be considered simultaneously for classification. D represents the standardised discriminant score for case i. Like APT, this score could represent the risk score or index. The mean of each discriminant function over all cases is zero and the standard deviation for D_i is 1. Just as D_i can be calculated for each case, a mean value of D_i can be calculated from each classification category. Membership into either group is then determined by whether or not the case's D score falls within a standard deviation of the group mean. Group means are known as centroids in reduced space, the spacing having been reduced from that of the k predictors to a single dimension, or discriminant function.

To assign cases into the three groups, either the two discriminant functions are used, or three classification equations are developed. In is simplest form, the basic classification equation for the jth group (j = 1,2....k) is

(2.25)

$$C_j = c_{j0} + c_{j1}X_1 + c_{j2}X_2 + \ldots\ldots + c_{jk}X_k$$

Testing of either classification equations rest on their ability to classify known cases from the sample, i.e., the proportion of cases classified correctly. If the classification rate is unsatisfactory, then either the estimation procedures are incorrect or the predictor variables inadequate for classification purposes, i.e., insufficient variance is explained by the independent variables.

One of the major advantages of classification indices, based on probability of membership, is their robustness to the violation of variable normality if violation is caused by skewness rather than outliers (Tabachnic, Fidell 1989). A major disadvantage of discriminant analysis is its assumption of linearity among all pairs of predictor factors.

Given that classification of cases is dependent on predictor factors, it then becomes important that the 'correct' variables are selected in constructing the discriminant functions. An advantage of this form of evaluation is the ability to construct composite variables of like factors through use of multifactor or principle components analysis. This procedure can ensure that independent variables are truly independent.

Many banks require the answers to standard personal questionnaires presented to applicants for credit. The results of these questions can be combined with other known physical factors related to the activity intended by the applicant. A study by Krause and Williams (1986) developed a behavioral model of farm families that included factors such as motivation, ability, and biographic variables. Their study suggested that personality variables, of both husbands and wives, were important in developing systems for use in evaluation of farm credit applications.

Variables used in the study included risk aversion (m), scientific orientation (m), manifest anxiety (m), anxiety score (w), authoritism (m), vocational interest (w), adaptability (m) scientific knowledge (w), aggressive conservatism (m) and unresolved rebellion (w), where (m) indicates for men and (w) indicates for women. All variables were significant. The diversity of potential variables used in a discriminant function analysis designed to construct 'discriminant risk scores' would appear to be unlimited. This supports the view that risk, and its various components, is a very complex multifactorial concept. Any attempt to quantify risk must take into account this fact.

2.4 RISK QUANTIFICATION

The underlying theme of all traditional methods of risk quantification is the acceptance of variance as the basic measure of risk and uncertainty. Much of the work done has centred on the decision process under conditions of risk and uncertainty, usually from a 'choice of alternatives' point of view. Using the EV framework in conjunction with either a linear or quadratic programming algorithm, to either maximise utility or minimise variance, concentrates solely on the decision problem. It takes variance for granted, as do beta models, as the quantification yardstick for risk.

As noted in previous sections, researchers have questioned the validity of variance as the primary surrogate measure of risk. Given the complexity of deriving higher moments from combinatorial distributions, and the difficulty of identifying and empirically estimating the parameters of individual distributions, researchers have sought to overcome the issue of risk quantification by ignoring it in favour of assuming that it forms part of the 'black box' behaviour of individual utility or preference functions, and that measuring utility therefore indirectly measures risk. This implies that risk is a qualitative concept based on personal belief, rather than a measurable quantitative concept (Anderson, Hardaker 1973).

Where the EV criteria method for activity selection has been refuted as a decision framework, it has been replaced by methods that rely on stochastic dominance or efficiency. SD relies on knowing the relevant cumulative density functions of payoffs and selecting that payoff, for example, whose CDF lies to the right of all others, i.e., first degree stochastic dominance (Hadar, Russell, 1969). The problem with this method is that the relevant CDFs need to be measured, and the analysis is still concerned with individual choice and preference. One danger with SD is its propensity to eliminate from consideration low variance portfolios, even though these portfolios also had low returns (Porter, Gaumnitz, 1972).

A danger with stochastic dominance as a decision criteria for farm activity selection is that it could support the kind of heuristics that people use under conditions of uncertainty. SD supports subjective income probability distributions very well and outperforms EV given a subjective probability distribution, but performs poorly in comparison to EV given an objective income probability distribution (Lee, Brown, Lovejoy, 1985). Given the results from this study, a banker might be more inclined to accept the more objective framework of EV rather than SD, since bankers have often been at the mercy of heuristic farmer forecasts of future income.

The issue addressed in this chapter is whether or not variance, in whatever form it is used, is a good surrogate risk measure. The decision problem is solved on the basis of risk, once risk is quantified. Utility and preference functions become usable according to the risks inherent in any activity. Given variance as a measure, the decision maker cannot make a decision unless he/she has a comparison.

What is required is a yardstick measure that all can simultaneously relate too, in terms of whether or not an activity or investment is risky or not risky, and know from the measure exactly how risky that activity is. An individual's preference ordering and risk aversion then applies in his decision in terms of how much risk he is prepared to accept, and he either accepts or rejects the investment on the basis of that risk measure.

Defining risk as the probability of loss or the probability of not achieving an expected outcome, rather than the expected dispersion or fluctuation about an expected outcome, immediately takes us away from any analysis of the X axis of any probability density distribution regardless of whether or not the distribution is subjective or objective. If probability is measured by relative frequency, objective or subjective, i.e., the Y axis of the distribution, then given a distribution, we are logically interested in the kurtosis and mode of that distribution relative to our expectation. If kurtosis is correlated with skewness then we are also interested in skewness.

Chapter Three

LINEAR PROBABILITY MODELLING

3.1 AN ALTERNATIVE APPROACH TO RISK MEASUREMENT

If risk is multifactorial and defined as the probability of loss or harm, and if variance analysis of the X axis of a net returns distribution is inappropriate as a risk measurement, then the Y axis or probability measurement is the only alternative. If an expected return has been identified, then an expected loss, or risk in not achieving that expectation, is that area to the left of the expectation under the distribution. The probability of the actual outcome falling below the expectation is the cumulative probability up to the expectation.

Given those factors that might influence the level of risk, or probability of loss, it is possible to estimate directly that level of risk, i.e., directly estimate the probability of an outcome being less than expected. If a minimum expected level of outcome is defined for any activity, below which loss is incurred, then the estimate of the probability of that defined outcome can be made, based on known historical outcomes. The resulting probability of a defined limit occurring will then constitute a measure of risk.

3.2 A LINEAR PROBABILITY MODEL - THE LOGIT MODEL

3.2.1 ORDINARY LEAST SQUARES AND THE LINEAR PROBABILITY MODEL

The 'logit' model, or loglinear analysis, is an extension of the multiway frequency analysis of the relationship between several categorical, or qualitative, and discrete variables. It has become an accepted framework for the analysis of a wide range of social, medical and economic issues (Berkson 1951, McFadden 1974). Alternatively known as regression on a dichotomous dummy dependent variable, the method utilises a weighted linear sum to predict the classification of cases into categories crosstabulated according to independent explanatory variables.

Assuming that a predetermined minimum threshold outcome for any activity or portfolio, below which loss or harm will occur, is defined as ϕ, consider the following simple model:

(3.1)

$$Y_i = \alpha_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots\ldots + \beta_n X_{ni} + u_i$$

Where X_{1i} ..... X_{ni} = n independent factors influencing risk and where $Y_i = 1$ if $i <= \phi$ (the threshold outcome level)

$Y_i = 0$ if $i > \phi$ (otherwise)

The above model is a linear probability model since $E(Y_i \mid X_{ni})$, the conditional expectation of $Y_i = 1$ given the X_is, can be interpreted as the conditional probability that the event will occur given X_{ni}, i.e., $\Pr(Y_i = 1 \mid X_{ni})$.

Assuming $E(u_i) = 0$ (to obtain unbiased estimators),

(3.2)

$$E(Y_i|X_{ni}) = \alpha_0 + \beta_{1i}X_{1i} + \beta_{2i}X_{2i} + \ldots\ldots + \beta_{ni}X_{NI} + u_i$$

Letting P_i equal the probability that $Y_i = 1$ (the threshold occurs) and $1 - P_i$ the probability that $Y_i = 0$ (the threshold does not occur), the variable Y_i has the distribution

Y_i	Probability
0	$1 - P_i$
1	P_i
	1

Therefore by the definition of mathematical expectation,

(3.3) $$E(Y_i) = 0(1 - P_i) + 1(P_i) = P_i$$

and can equate to obtain

(3.4)

$$E(Y_i|X_{ni}) = \alpha_0 + \beta_{1i}X_{1i} + \ldots\ldots + \beta_{ni}X_{ni} + u_i = P_i$$

The conditional expectation of Y_i can be interpreted as the conditional probability of Y_i (Gujarati 1988).

Since the probability P_i must lie between 0 and 1, the constraint that $0 <= E(Y_i|X_{ni}) => 1$, i.e., the conditional probability, must lie between 0 and 1.

It would appear that the above model could now be estimated by the standard OLS method. Unfortunately, even though OLS assumes the disturbances (u_is) to be normally distributed, the assumption is no longer tenable because like Y_i, u_i takes on only two values.

(3.5)

$$u_i = Y_i - \alpha_0 - \beta_{1i}X_{1i} - \ldots\ldots - \beta_{ni}X_{ni}$$

$$\textit{when} \quad Y_i = 1 \quad u_i = 1 - \alpha_0 - \beta_{1i}X_{1i} - \ldots\ldots - \beta_{ni}X_{ni}$$

$$\textit{when} \quad Y_i = 0 \quad u_i = -\alpha_0 - \beta_{1i}X_{1i} - \ldots\ldots - \beta_{ni}X_{ni}$$

U_i follows a binomial distribution, and cannot be assumed normally distributed (Gujarati 1988).

In addition, even if $E(u_i) = 0$ and $E(u_i u_j) = 0$ for i not equal to j, i.e., no serial correlation, the disturbance terms are heteroscedastic. To see this, following from the previous distribution, the u_is have the probability distribution

u_i	Probability
$- \alpha_0 - \beta_{ni}X_{ni}$	$1 - P_i$
$1 - \alpha_0 - \beta_{ni}X_{ni}$	P_i
	1

By definition

(3.6) $$Var\ (u_i) = E[u_i - E(u_i)]^2 \quad = E(u_i^2)$$

$$for \quad E(u_i) = 0 \quad by\ assumption$$

and using the probability distribution of u_i, we obtain (3.7)

$$Var\ (u_i) = E(u_i^2) = (-\alpha_0 - \beta_{ni}X_{ni})^2(1 - P_i) + (1 - \alpha_0 - \beta_{ni}X_{ni})^2(P_i)$$

$$= (-\alpha_0 - \beta_{ni}X_{ni})^2(1 - \alpha_0 - \beta_{ni}X_{ni})$$

$$+ (1 - \alpha_) - \beta_{ni}X_{ni})^2(\alpha_0 + \beta_{ni}X_{ni})$$

$$= (\alpha_0 + \beta_{ni}X_{ni})(1 - \alpha_0 - \beta_{ni}X_{ni})4$$

$$with \quad Var(u_i) = E(Y_i|X_{ni})[1 - E(Y_i|X_{ni})]$$

$$= P_i(1 - P_i)$$

The variance of u_i is heteroscedastic because it depends on the conditional expectation of Y, which depends on the values taken by the X_is. Thus ultimately the variance of u_i depends on the X_is and is thus not homoscedastic (Gujarati 1988). Therefore, under heteroscedasticity, OLS estimators will not be efficient, although weighting of variables would handle the heteroscedasticity problem.

A further problem is that there is no guarantee that Y_i, the estimators of $E(Y_i|X_i)$ will lie between 0 and 1. This can be achieved by resorting to restricted least squares or mathematical programming techniques, but these are complicated procedures.

The most basic problem with OLS and linear probability modelling is the fact that LPM does not appear logically attractive. It assumes that $P_i = E(Y_i = 1|X_{ni})$ increases linearly with X_{ni}, i.e., the marginal or incremental effect of the X_{ni}s remains constant. What is therefore needed is a probability model that has two features: (a) As X_{ni} increases, $P_i = E(Y = 1|X_{ni})$ increases but never steps outside the 0 to 1 interval, and (b) the relationship between P_i and the X_{ni}s is nonlinear, where the approach to zero is slower and slower as X_{ni} becomes small, and approaches 1 at slower and slower rates as X_{ni} becomes large.

In effect, such a model would reflect the decreasing risk of an individual activity, decreasing until negligible as expected net returns for that activity grow larger and larger. Should the historical nature of an agricultural enterprise indicate a 'safe bet' with regard to relatively high net returns, then intuitively the risks associated with that activity should be small, i.e., the probability of that activity's actual outcome occurring at the predetermined threshold of loss should be extremely low.

3.2.2 THE LOGISTIC FUNCTION AND LOGIT MODELLING

The cumulative distribution function, which describes the probability of any expected outcome, is sigmoid shaped. Its slope at any point is determined by the kurtosis and skewness of its underlying probability density function. Kurtosis regulates the rate at which the CDF ascends from 0 to 1, and skewness determines the rate of convergence with either 0 or 1 probability, at either tail.

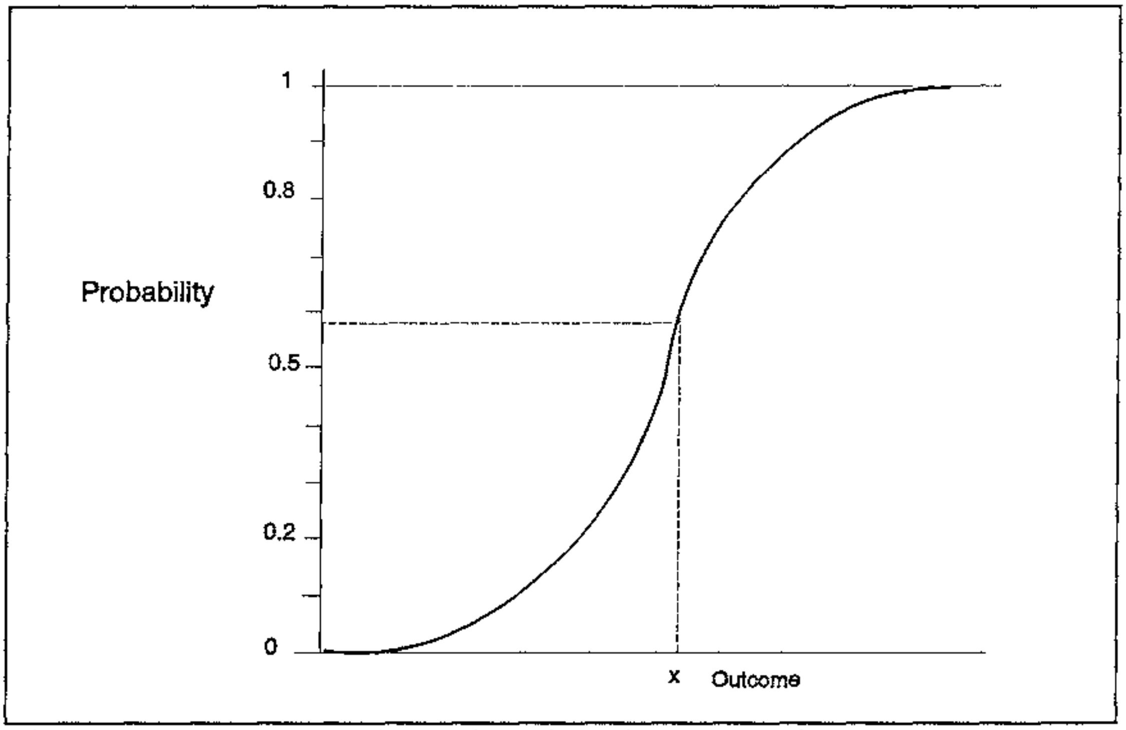


Figure 3.1 Cumulative Distribution function

Figure 3.1 indicates such a sigmoid function. All variable distributions have their associated CDF, and although all CDFs are sigmoid, every CDF is empirically unique according to the kurtosis and skewness of its underlying distribution. This raises the practical problem of which CDF model to use in an analysis. Historically, the CDF forms commonly chosen to describe dichotomous qualitative response models are the logistic and normal CDF, the former giving rise to the 'logit' and the latter the 'probit' or 'normit' model (Gujarati 1988).

With an underlying normal, or lognormal distribution, zero on the X axis will always equate to 0.5 on the probability scale. This is because zero will be the mean, and mode, of the distribution, given that the data is positively and negatively symmetrical about zero. In general, the mean of a distribution will always equate to 0.5 on the probability scale of its CDF. This basic fact is important in the context of a decision rule that uses CDF probability risk measures (Gujarati 1988).

The general form of the logistic function is

$$P_i = \frac{1}{1 + e^{-z_i}} \tag{3.8}$$

where e is the base of the natural logarithm, and z_i represents the linear sum of the independent variables times their coefficients, i.e.;

(3.9)
$$z_i = \alpha_0 + \beta_{1i}X_{1i} + \ldots\ldots\ldots + \beta_{ni}X_{ni} + u_i$$

Thus the full logit model takes the form

(3.10)
$$P_i = E(Y = 1|X_{ni}) = \frac{1}{1 + e^{-(\alpha_0 + \beta_{1i}X_{1i} + \ldots + \beta_{ni}X_{ni})}}$$

As z_i ranges from negative infinity to positive infinity, P_i ranges between 0 and 1, and P_i is nonlinearly related to z_i. But although these two preconditions are satisfied, an estimation problem exists in that P_i is not only nonlinear in X_{ni} but in the β_{ni}s as well, therefore necessitating the transformation of the model, as follows.

If P_i is the probability of the threshold ϕ occurring, then the probability of it not occurring is

(3.11)
$$1 - P_i = \frac{1}{1 + e^{z_i}}$$

and therefore we can write

(3.12) $$\frac{P_i}{1 - P_i} = \frac{1 + e^{z_i}}{1 + e^{-z_i}} = e^{z_i}$$

Where P_i / $(1 - P_i)$ is now the 'odds ratio' in favour of the threshold being met, ie., the ratio of the probability of the actual outcome equalling ϕ to the probability that ϕ will not actually occur. Taking the natural logarithm of the above equation linearises the relationship between dependent and independent variables.

(3.13)

$$L_i = \ln(\frac{P_i}{1 - P_i}) = Z_i = \alpha_0 + \beta_{1i}X_{1i} + \dots + \beta_{ni}X_{ni} + u_i$$

Where L_i is the log of the odds ratio, and is in fact the 'logit'. As P goes from 0 to 1, the logit L goes from negative to positive infinity. The logits are not bounded between 0 and 1, and although L is linear in X, the probabilities are not (Gujarati 1988).

The coefficients of the model indicate the change in L for a unit change in X, i.e., the effect a unit increase of X has on the log-odds of our actual outcome being equal to ϕ. To derive the probability P_i of the actual outcome = ϕ (the threshold of loss incurred), and obtaining the 'risk index' or 'credit rating', we need simply solve the logistic function

(3.14)
$$P_i = \frac{1}{1 + e^{-L_i}}$$

3.2.3 ESTIMATION OF THE LOGIT COEFFICIENTS

Given the model

(3.15)
$$L_i = \ln\left(\frac{P_i}{1 - Pi}\right) = \alpha_0 + \beta_i X_i + u_i$$

and putting the values 1 and 0, according to whether or not our observed outcomes fall above or below ϕ, directly into the logit L_i, we obtain: $L_i = \ln(1/0)$ if our observation is less than or equal to ϕ, and $L_i = \ln(0/1)$ if our observation is greater than ϕ. These expressions are meaningless and enforce use of the maximum-likelihood method to estimate the parameters, particularly if data is at the micro or individual level (Gujarati 1988).

Logit analysis transforms the observed individual data into the relative frequency of individual cases falling into either category 1 or 0, according to the independent factors. If X_i

represents farm area for example, N_i the number of observations in each frequency class of X_i, and n_i the number of observations that fall into category 1 (less than or equal to ϕ) at each class level of X_i, i.e., corresponding to each level of area X_i, there are N_i farmers, n_i among whom are classified into category 1, then if N_i is large, the relative frequency n_i/N_i can be used as an estimate of the true P_i. Fitting according to the logit equation will estimate the true logit L_i well if N_i is large (Gujarati 1988).

As with LPM, the disturbance terms in logit analysis are also heteroscedastic. It can be shown that if N_i is fairly large and if each observation in a given X_i class is distributed independently as a binomial variable, then u_i follows a normal distribution with zero mean but variance equal to $1/[N_iP_i(1-P_i)]$.

$$u_i \sim N[\ 0,\ \frac{1}{N_i P_i (1 - Pi)}] \tag{3.16}$$

This makes necessary the use of weighted maximum likelihood estimators, where the standard deviation of u_i is used as the weight, i.e.,

$$\hat{\sigma}^2 = \frac{1}{N_i \hat{P}_i (1 - \hat{P}_i)} \tag{3.17}$$

since $P_i = n_i/N_i$, L_i can alternatively be expressed as $L_i = \ln n_i/(N_i - n_i)$. To avoid P_i taking the value 0 or 1, in practice L_i is measured as:

(3.18)
$$\hat{L}_i = \ln \frac{(n_i + \frac{1}{2})}{(N_i - n_i + \frac{1}{2})} = \ln \frac{(\hat{P}_i + \frac{1}{2}N_i)}{(1 - \hat{P}_i + \frac{1}{2}N_i)}$$

3.2.4 INTERACTION EFFECTS

As well as continuous variables, dichotomous variables are also valuable as independent variables in the logistic function. They can be used as dummy variables, in the same way that they are used in multiple regression, for sorting out the effect on the logit of different categories of exogenous factors. In fact it is possible to have the entire logit function dichotomous or trichotomous in both dependent and independent variables (Goodman 1972).

Consider the following model:

(3.19)
$$L_i = \alpha_0 + \beta_{1i}X_{1i} + \beta_{2i}X_{2i} + \lambda_{1i}D_{1i} + \lambda_{2i}D_{2i} + u_i$$

Where X_1 and X_2 are any exogenous factors. D_1 and D_2 are dummy variables, both dichotomous yes or no (1 or 0). A fully saturated model, totally predicting classification into $Y_i = 1$ or 0, could contain not only main effects but also two, three and four way effects, ie.,

Main effects	X_1 X_2 D_1 D_2
Two way	X_1X_2 X_1D_1 X_1D_2 X_2D_1 X_2D_2 D_1D_2
Three way	$X_1X_2D_1$ $X_1X_2D_2$ $X_1D_1D_2$ $X_2D_1D_2$
Four way	$X_1X_2D_1D_2$

Each interaction variable coefficient describes the impact on L_i of the specific qualitative characteristics described by the dummy variable. For example, if D_1 represents a provincial region in New Zealand, and D_2 a topographical class of farm, say Northland for D_1, $D_1 = 1$ if yes, 0 if no, and North Island Hill country for D_2, $D_2 = 1$ if yes, 0 if no, then the interaction coefficient describes the adjustment effect on L_i, of hill country farming in Northland, i.e., there exist unique risks in that type of farm in that region.

3.3 THE DECISION FRAMEWORK WITHIN LOGIT ANALYSIS

Given that the logit model

$$P_i = Pr[E(Y_i = 1|Z_i)] = \frac{1}{1 + e^{-Z_i}} \tag{3.20}$$

derives the probability P_i of $Y_i = 1$ given Z_i, where $Y_i = 1$ if $i <= \phi$, then changes in Z_i shift the logistic function according to P_i.

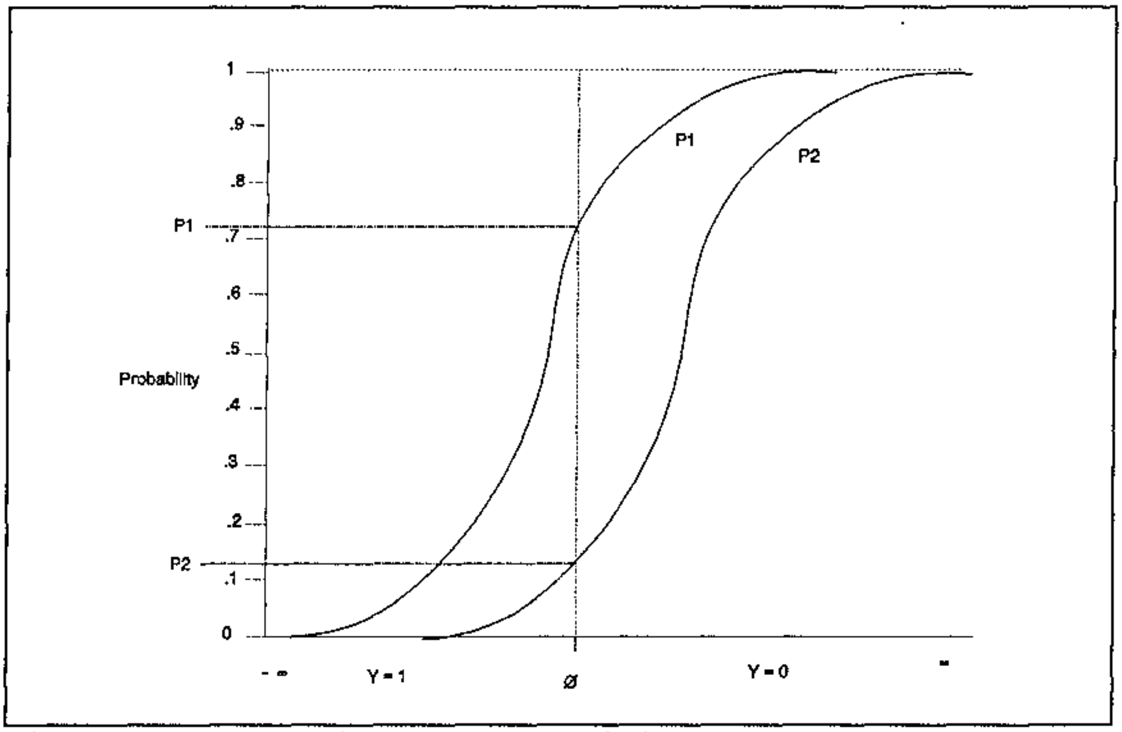


Figure 3.2 Relative CDF Probability

Figure 3.2 displays two cumulative density functions P1 and P2, which could be described as representing either two individuals with various mixes of activity or product, or two individual activities. The logit function has derived probabilities for both, i.e., P1 and P2 (0.73 and 0.13 respectively), representing the probabilities of each activity being equal to or less than ϕ, a threshold level of risk, i.e., where $Y_i = 1$.

According to the probability measures, one would prefer to select P2, given that its measure of risk is less than that of P1. By utilising the risk measurement, one is immediately distinguishing among distributions and making selections according to the underlying distributions of each option.

Let P1 represent G(x) and P2 represent F(x), both CDFs for the distributions f(x) and g(x). The question is whether or not F(x) would be preferred over G(x) by all decision makers regardless of their preference functions. The answer determines the generality of acceptance of P_i as a measurement of risk for a yes/no decision on an independent activity or choosing between two or more options. Will all decision makers, regardless of their preference functions, select F(x) over G(x), and will all decision makers, with a predefined risk threshold level, accept P_i if P_i exceeds their threshold of acceptance ?

The stochastic dominance problem of finding necessary and sufficient conditions on cumulative distributions G(x) and

F(x), such that F(x) will be preferred or indifferent to G(x) by all agents in a particular group of agents, has been solved by Hadar and Russell (1969), Hanoch and Levy (1969), Meyer (1975,1976), and others.

Assuming all preferences of a decision maker are represented by an expected utility function u(x) which is increasing and twice differentiable, groups are then described in terms of the properties of their utility functions. Meyer (1975) does not find this method of defining groups very convenient, because u(x) is not a unique representation of a given set of preferences. To restrict the group of agents under consideration by restricting u(x), the restrictions used on u(x) must also be met by all positive linear transformations of u(x) because any positive linear transformation of u(x) also represents those same preferences (Meyer 1975). Generally, stochastic dominance places restrictions on u(x) by specifying the signs of the second and third derivatives of u(x). Because many possible restrictions on u(x) do not define groups of decision makers, since membership of a group depends on the particular representation of preferences being used, Meyer prefers to place restrictions on r(x), where r(x) is defined as

$$r(x) = \frac{-u''(x)}{u'(x)} \tag{3.21}$$

i.e., $r(x)$ represents decision maker's preferences uniquely (Pratt 1964). Restrictions on $r(x)$ correspond directly to restrictions on preferences, by describing groups of decision makers with $U(r_1(x), r_2(x))$ being the set of decision makers with preferences $r(x)$ satisfying

(3.22)
$$r_1(x) \leq r(x) \leq r_2(x) \qquad \forall x$$

or alternatively

(3.23)
$$r_1(x) \leq -\frac{u^{-u''(x)}}{u'(x)} \leq r_2(x) \qquad \forall x$$

for given functions $r_1(x)$ and $r_2(x)$. Pratt and Arrow interpret $r(x)$ as a measure of a decision maker's absolute aversion to risk. Thus restrictions on $r(x)$ can be viewed as an upper and lower bound on the degree of risk aversion for the decision makers in the set under consideration (Meyer 1975).

The expected utility assumption states that $F(x)$ is preferred or indifferent to $G(x)$ by an agent with utility $u(x)$ if and only if

(3.24)
$$\int_0^1 u(x)\,dF(x) \geq \int_0^1 u(x)\,dG(x)$$

or equivalently

$$\int_0^1 [G(x) - F(x)]\,u'(x)\,dx \geq 0 \tag{3.25}$$

For any two functions $r_1(x)$ and $r_2(x)$ the necessary and sufficient conditions on F(x) and G(x), for F(x) to be preferred or indifferent to G(x) by all decision makers in the class $U(r_1(x), r_2(x))$, are:

1. F(x) is preferred to G(x) by decision makers in the set

$$U(-\infty, +\infty) \quad if$$

$$[G(x) - F(x)] \geq 0 \qquad \forall\, x \subset [0, 1] \tag{3.26}$$

2. F(x) is preferred to G(x) by decision makers in the set

$$U(0, +\infty) \quad if$$

$$\int_0^y [G(x) - F(x)]\,dx \geq 0 \qquad \forall x \subset [0, 1] \tag{3.27}$$

The above two conditions on F(x) and G(x) are defined respectively as first and second-degree stochastic dominance of F(x) over G(x), where set 1 places no restrictions on preferences other than assuming u(x) is increasing and twice differentiable. The group in set 2 is important because it contains the set of all risk averse decision makers. $r_1(0)$ defines the risk threshold level below which decision makers are not prepared to accept risk.

Where the level of risk averseness can be defined by a function, i.e., $r_1(x)$ is any function, then given

$$r_1(x) = \frac{-u''(x)}{u'(x)} \tag{3.28}$$

3. F(x) is preferred to G(x) by agents in the set

$$U(r_1(x),\ +\infty) \quad if$$

$$\int_0^y [G(x) - F(x)]\, u_1'(x)\, dx \geq 0 \qquad \forall x \subset [0,1] \tag{3.29}$$

i.e., the condition that applies when only a lower bound on the risk aversion of a group of decision makers is specified.

Where only an upper bound on the risk aversion of a group is specified, ie., where r_2 is any function, then

4. F(x) is preferred to G(x) by agents in the set

$$U(-\infty,\ r_2(x)) \quad if$$

(3.30)
$$\int_y^1 [G(x) - F(x)]\, u_2'(x)\, dx \geq 0 \qquad \forall x \subset [0,1]$$

The situation where both r_1 and r_2 are specified lower and upper bound functions, is much more difficult to solve. Meyer suggests that since r(x) represents a given set of preferences the problem is solved for a given F(x), G(x), $r_1(x)$ and $r_2(x)$ by checking if the expected utility from F(x) is greater than G(x). His solution failed to display a closed form solution, and is in the form of a rule for calculating the solution in an applied way.

Maximising

(3.31)
$$-\int_0^1 [G(x) - F(x)]\, u'(x)\, dx$$

subject on

$$r_1(x) \leq [\frac{-u_0''(x)}{u_0'(x)}] \leq r_2(x) \quad with \quad u'(0) = 1$$

which is given by

(3.32)

$$\frac{-u_0''(x)}{u_0'(x)} = \begin{cases} r_1(x) & \text{if} \quad \int_x^1 [G(y) - F(y)]\, u_0'(y)\, dy < 0 \\ \\ r_2(x) & \text{if} \quad \int_x^1 [G(y) - F(y)]\, u_0'(y)\, dy \geq 0 \end{cases}$$

In general, theory would indicate that F(x) will be preferred to G(x) by all agents in the set $U(r_1(x), r_2(x))$. Therefore P_i as a measure of risk will generally be identified as discriminating between a high or a low risk option. Where risk aversion can be identified in terms of a P_i level for $r_1(x)$, i.e., a P_i lower bound, then any P_i greater than $r_1(x)$ will be accepted as possible options. Those options with P_i less than $r_1(x)$ will not be possible options, according to stochastic dominance conditions.

3.4 LOGITS OR PROBITS ?

In comparing the relative merits of utilising either the logistic function or integrated normal function for deriving P_i, a quick look at the functional forms of both will indicate that the functional simplicity of the logistic function is appealing.

The logistic function can be specified as:

$$y = \frac{1}{1 + e^{-(\alpha + \beta X)}}$$

(3.33)

$$logit(y) = \ln\frac{y}{1 - y} = \alpha + \beta x$$

The integrated normal function can be specified as:

$$y = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha + \beta X - 5} e^{(-t^2)/2}\, dt$$

(3.34)

$$probit(y) = normal\ deviate + 5 = \alpha + \beta X$$

The logistic curve is distinctly easier to fit than the integrated normal curve, and is particularly suited to binomially distributed variates (Berkson 1951). This arises from the fact that if the variate p is logistic, the first derivative is pq. Since the weight in fitting procedures is reciprocal to the binomial variance pq, these will frequently cancel. Thus the normal equations for MLE of the logistic function are:

$$\sum n(\, p - \hat{p}\,) = 0$$

(3.35)

$$\sum n(\, px - \hat{p}x\,) = 0$$

It is seen that the coefficients are dependent only on n, the sample size.

The coefficients of the normal equations for the integrated normal curve contain not only z, the ordinate of the normal curve, but pq, the values to be estimated.

$$\sum \frac{n\hat{z}}{\hat{p}\hat{q}}(\, p - \hat{p}\,) = 0$$

(3.36)

$$\sum \frac{n\hat{z}}{\hat{p}\hat{q}}(px - \hat{p}x\,) = 0$$

The relative precision of both minimum X^2 and maximum likelihood estimates of regression coefficients, appear to give the logistic function a definite advantage over the integrated normal curve as a statistical instrument (Berkson 1951). In addition the logistic estimates - both the X^2 and likelihood estimates - besides being asymptotically efficient are also statistically sufficient, while the estimates of the parameters of the integrated normal curve are not sufficient (Berkson 1951).

Pictorially both the logistic and integrated normal curves are strikingly similar, so much so that they are practically superposable, but their respective equations make one easier to handle than the other when solving for P_i, the probability risk measure, in terms of L_i, the logit, within the logistic function.

Given the assumption of non-linear relationships, and the estimation ease and generality of the logistic curve, it will is used in this study as the basic sigmoid curve representing the cumulative distribution of net cash returns for individual farm production units, from which probability measures of risk will be estimated. Using this technique assumes that the combinatorial cumulative density function representing all agricultural net cash returns, is logistically sigmoid shaped.

Chapter Four

DATA DESCRIPTION

4.1 SURVEY SAMPLE

Data utilised in the construction of the logit model was provided by the New Zealand Meat and Wool Boards' Economic Service, from annual sheep and beef farm survey data gathered for the years 1984/85 to 1988/89 inclusive. Their survey uses actual farm account information for revenues, expenditures and balance sheet data, and is historical in nature. Physical data are collected separately for each farm in the survey and data is compiled on a per farm basis, regardless of ownership structure.

The annual sample units are selected randomly and the sample made large enough to avoid bias from unit abnormality. Each and every qualifying farm unit has had equal chance of selection. Qualifying is defined as having wintered at least 750 sheep or their equivalent sheep plus cattle stock units, must be privately operated as an independent ordinary commercial farm unit, and at least 80 per cent of the farm revenue derived from sheep or sheep plus beef cattle (except in the case of mixed finishing farms located in Canterbury) (NZ Sheep & Beef Farm survey 1988-89).

The population consists of a comprehensive list of sheep and beef farm owners from which a random sample is drawn and divided or stratified according to geographical area, flock size and farm class. In addition, use of variable sampling fractions is made to reduce random sampling error. This means that the population is divided into homogeneous groups of strata from which a random sample is taken. This ensures that groups within the population are adequately represented. Stratification has the objective of spreading the total annual sample of approximately 530 farms over the main sheep and beef farming districts, and flock sizes within regions using random selection proportionate to the regional and flock size distributions of sheep and beef farms.

All Crown properties and farms with less than 750 stock units are excluded before sampling. This excludes about 7 per cent of sheep and beef farms, reducing the population by over twelve thousand flocks. The annual sample of 530 represents a qualifying population of 21,300 sheep and beef farms, or 2.5 percent of the population.

4.2 SAMPLE UTILISATION

Given that the objective of this thesis is to discover the probability of relationships between nominal net farm cash incomes and factors likely to influence net cash incomes, including nominal prices received for farm products, it was decided that the five independent annual farm surveys be combined into one sample representing the five year time frame.

The nature of logit model and coefficient estimation need not include time adjustment parameters, as the process involves the prediction of the probability of membership of cases to the categories of a dichotomous dependent variable according to the crosstabulation of that dependent variable with other discrete independent variables. This is not a regression exercise involving the exact prediction of a dependent outcome, nor the direct relationship between dependent and independent variables.

Table 4.1 displays the annual sample sizes for each accounting year where an accounting year varies according to the individual farms' balance date in any year. A physical production year, measuring stock transactions, is always June ending. The sample percentage of farm population for the years 84/85 to 87/88 inclusive are based on a population of 22,000 farms, and for 88/89 on a population of 21,300. The final combined sample size percentage of 11.73 is based on 22,000 farms.

Table 4.1 ANNUAL SAMPLE SIZES

ACCOUNTING YEAR	SAMPLE SIZE	% OF POP.
1984/85	520	2.36
1985/86	510	2.32
1986/87	503	2.27
1987/88	513	2.33
1988/89	535	2.51
TOTAL SAMPLE	2581	11.73

4.3 VARIABLE DEFINITIONS

Appendix I presents the list of variables requested from the NZ Meat & Wool Boards' Economic Service. Variables selected for and utilised in the modelling process were either constructed from the Appendix I list, or used directly. The following variables were selected for model testing according to their possible influence on net farm returns and their likely impact on the probability of a farm incurring negative returns after a years' trading. Names in brackets are the actual model variable names used in the analysis.

CLASS (CLASS)

The topographical farm class classifications, as used by both the NZ Statistics Department and Meat & Wool Boards Economic Service, includes classifications 1 through 8.

1. SOUTH ISLAND HIGH COUNTRY: Extensive high country run properties carrying fine wool sheep with wool as the predominant source of income. Situated mainly in Marlborough, Canterbury and Otago.

2. SOUTH ISLAND HILL COUNTRY: Rolling to steep hill country with mainly finer wooled sheep and carrying about three stock units per hectare. Wool and stock sales the main source of revenue. Situated mainly in Canterbury.

3. NORTH ISLAND HARD HILL COUNTRY: Steep hill country with about eight stock units per hectare and a ratio of 12 sheep per cattle beast and approximately 75 percent of income from sheep production. Situated mainly on the East and West coasts and the central plateau of the North Island.

4. NORTH ISLAND HILL COUNTRY: Easier rolling hill country with mainly crossbred wool type sheep averaging 10 stock units per hectare and a ratio of 11 sheep per cattle beast. Revenue mainly from forward and prime sale stock. These properties are situated throughout the North Island.

5. NORTH ISLAND INTENSIVE FINISHING FARMS: High production grassland farms located on either flat or flat/rolling hill country carrying 12 stock units per hectare with a ratio of 11 sheep per cattle beast. Situated mainly in South Auckland, Taranaki, King Country, Hawkes Bay and Wellington.

6. SOUTH ISLAND FINISHING/BREEDING FARMS: Extensive type of easy country finishing farms breeding own replacements accompanied by moderate levels of cash and fodder cropping. Mainly found around Canterbury, Otago and parts of Southland.

7. SOUTH ISLAND INTENSIVE FINISHING FARMS: High production flat grassland farms carrying 12 stock units per hectare on average, with some cash cropping. Found mainly in Southland and parts of Otago.

8. SOUTH ISLAND MIXED FINISHING FARMS: Mainly found on intensive Canterbury farms with a high proportion of income from seed and cash crop production, as well as finishing stock.

REGION (LOC.)

The regional location of farms is specified according to the counties and districts as they were during 1986. Figure 4.1 indicates the major regions, and Appendix II the counties and districts within each region. There are no farms from the West Coast region of the South Island in the sample, and is due to the limited number of qualifying properties found there.

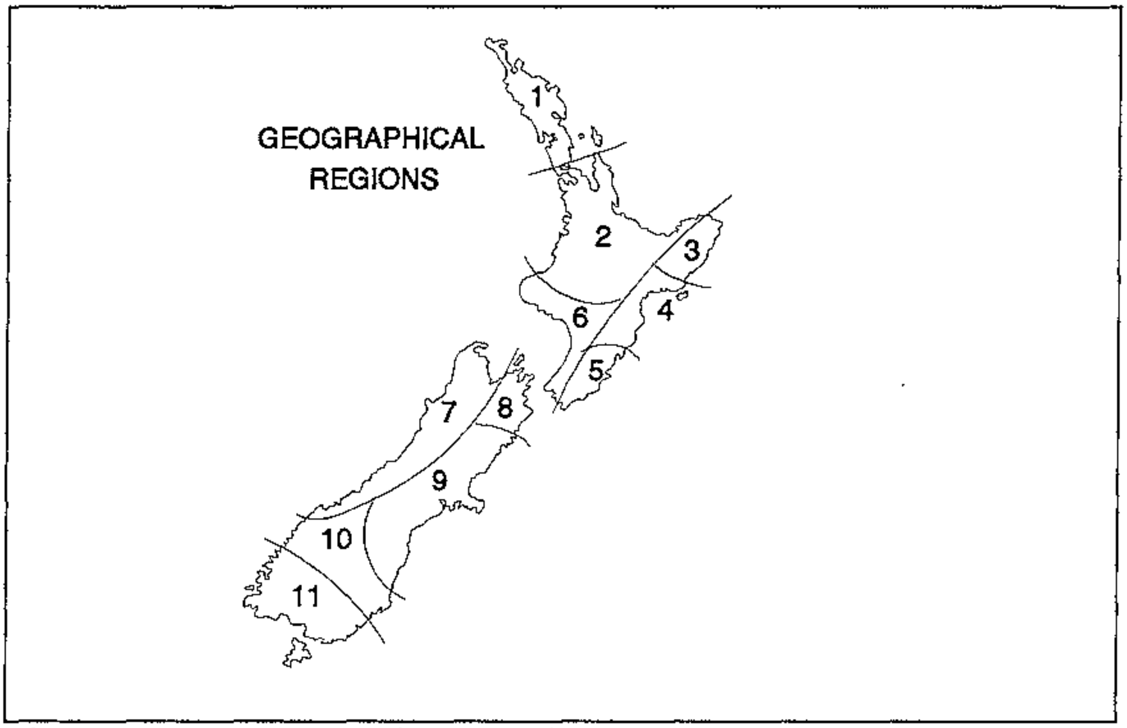


Figure 4.1 Geographical Regions

AREA (AREA)

This is effective area in hectares, of productive land. Excludes bush reserves, buildings and raceways or retired land.

NET INCOME (NTINC)

Net Income is defined as the CASH returns net of all cash expenditures, according to the definition outlined in chapter one. Total income comprises all cash revenues from the sale of all production minus all stock purchases. No allowance is made for changes in livestock values, as these are intangible value adjustments. Similarly, total cash expenses comprises all farm working expenses, capital purchases, personal drawings, taxation payments and all mortgage payments. Depreciation is considered an intangible non cash entity and is not included as an expense. A negative net income in the context of this thesis, indicates an inability to service all private and farm expenses, as well as lending, from actual on farm production.

WOOL PRICE (WOOL)

This is the average price at auction of all wool type sales, i.e., the total gross wool revenue divided by the total kilograms of wool sold, and is defined as the average gross per kilogram dollar price of wool at auction.

LAMB PRICE (LAMB)

This is the average per head price received for all types of lamb sales, i.e., the sum of prime, store and live whether, ram, ewe and cryptoid lamb sale incomes divided by the sum of prime, store and live sale numbers.

EWE PRICE (EWE)

The average per head price received for all types of mutton, store, cast for age and two tooth ewe. As above, the sum of all cash returns divided by the total numbers sold.

BEEF PRICE (BEEF)

The average per head average price received for all ages of steer cattle sales. Constructed as all prime and all age store sales divided by total numbers sold.

HEIFER PRICE (HEIF)

As above, the per head average price received for all ages of heifer cattle sales. Constructed as above.

TOTAL EXPENSES (TTEXP)

Defined as the total cash expenses of net cash incomes. It is the sum of all farm and personal cash expenditures, including capital purchases and excluding depreciation expenses.

EQUITY PERCENT (EQ %)

Defined as the percentage of farmer balance sheet equity of total asset values. Land, buildings and livestock are valued at market value. Plant is shown at book value. This equity includes investments and off-farm assets valued at cost.

FERTILISER (FT/HA)

The tonne per effective hectare of all fertiliser applications, including nitrogenous fertilisers.

STOCK UNITS (SU/HA)

Defined as the per effective hectare stock unit ewe equivalents, present at the 30th June, of only those livestock classes used in this analysis. Excluded from total stock units are; wethers, weaner bulls, unmated cows and heifers, deer and goats. Total stock units were calculated using the following conversion ratios:

Ewes	1 SU
2t Ewes	1 SU
Hoggets	0.7 SU
Rams	1.1 SU
Mated Cows and Heifers	6 SU
Yearling Cattle	4 SU
2 year Cattle	4.5 SU
3 year Cattle	5 SU
Bull Beef	6.5 SU
Breeding bulls	5.5 SU

LAMB SALE PERCENT (LB %)

Defined as the lambing percentage net of lamb deaths as at 30th June. Calculated as the sum of all lamb sales as a percentage of the open breeding numbers of ewes on hand at 30th June. Although not a lambing percentage figure because it includes the sale of lambs purchased, this method appealed because it included allowance for the deaths of lambs prior to balance date.

WOOL PRODUCTION (WL/HA)

Defined as the total kilograms of wool sold per effective hectare per year.

4.4 VARIABLE DESCRIPTIONS

Table 4.2 reports the basic statistics of the variables utilised in the building of the model.

Table 4.2 **STATISTICAL DESCRIPTIONS**

	MEAN	MODE	STD	SKEW	KURT	RANGE	MIS
NTINC	-17521	-44480	51256	-1.24	9.746	752899	0
WOOL	4.23	2.381	1.163	5.223	40.591	15.697	1
LAMB	17.10	12.00	5.095	0.695	3.496	59.974	61
EWE	12.12	6.00	5.490	2.593	17.623	66.891	50
BEEF	490.63	350.00	139.7	0.360	0.251	1012	850
HEIF	364.79	450.00	99.29	1.586	10.916	1345	880
TTEXP	153487	72855	109215	2.326	8.238	942134	0
FT/HA	0.195	0.250	0.172	1.181	1.995	1.187	557
SU/HA	9.620	9.489	3.790	-.480	0.225	22.826	0
LB %	97.251	100	31.30	2.013	22.045	179	9
WL/HA	38.80	12.786	20.71	0.548	0.345	112.74	0
AREA	967.15	245	2631	6.414	50.818	30113	0
EQ %	72.7	100.00	24.53	-1.64	4.051	200.6	0
CLASS	NA	NA	NA	NA	NA	NA	0
LOC.	NA	NA	NA	NA	NA	NA	0

MEAN = sample Mean, MODE = Sample mode,

STD = Sample standard deviation, calculated as

$$\sqrt{\frac{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}{N - 1}} \tag{4.1}$$

SKEW = Sample distribution skewness, calculated as

$$\frac{E[(Y_i - \overline{Y})^3]}{\sigma^3} \tag{4.2}$$

KURT = Sample distribution kurtosis, calculated as

$$\frac{E[(Y_i - \overline{Y})^4]}{\sigma^4} - 3 \tag{4.3}$$

RANGE = Sample data range, calculated as maximum minus minimum.

MIS = Number of missing observations.

All distributions are skewed with the exception of stock units per hectare, where both the mean and mode are similar. Sample distributions of all variables above are reported in Appendix III. Because logit analysis is performed on case number observations within crosstabulation cells, distributions are displayed in Appendix III as sample number observations rather than percentages. This information is useful for constructing the category ranges within the logit model independent variables, where quantitative variables need to be transformed into qualitative variables.

NTINC (Net Income), TTEXP (Total cash Expenses) are expressed in dollar nominal values, WOOL, LAMB, EWE, BEEF & HEIF are the average per head prices received in dollar amounts, FT/HA (fertiliser applications per hectare is expressed in tonnes per hectare and AREA in hectares.

Table 4.3 reports the sample distribution across both farm class categories and regions. All basic statistics refer to the combined annual surveys, i.e., means are the mean five year price.

Table 4.3 CLASS & REGION STRATA

CLASS	NO.	REGION	NO.	REGION	NO.
1	142	1	86	9	603
2	206	2	422	10	136
3	371	3	147	11	193
4	772	4	456		
5	429	5	141		
6	362	6	320		
7	187	7	0		
8	112	8	77		
TOTAL	2581				2581

4.5 VARIABLE SELECTION

As was stated earlier in this chapter, variables were selected on the basis of their possible influence on net cash incomes. More importantly, in the context of this study, variables were also selected according to whether or not they were considered to be 'risky'. This was a qualitative assessment based not only on their unpredictable fluctuating nature but also on whether or not they constituted a resource constraint that hinders the ability of the farmer to react to unforeseen circumstances.

Product prices are a significant 'risky' component of farm production. From both the financiers and farmers points of view, a price forecast upon which facility for seasonal finance has been granted does indeed constitute risk, as a large proportion of production investment is made, or expenditure incurred, well before the actual price is realised. The potential actual end of year net cash income is determined by the control or management of farm expenses prior to harvest and sale of product, but the realised actual end of year net cash income is determined by the relative uncontrollable nature of actual product prices received. For this reason both product prices and total farmer expenses should be selected for modelling, regardless of the historical statistical relationships between net incomes, expenses and product prices.

As part of a preliminary investigation into the nature of the variables under consideration for the logit model building process, a correlation matrix of all variables was developed. This matrix, Table 4.4, serves as a forewarning of any possible interaction effects to be considered in the model. Because NTINC is not a true regression type dependent variable, in the sense that a statistical relationship with independent factors is not being used to estimate actual NTINC, correlations should not be interpreted as a variable selection criteria.

Table 4.4 indicates the potential interaction terms that will require testing in the modelling process.

Table 4.4 VARIABLE CORRELATIONS

	NTINC	**WOOL**	**LAMB**	**EWE**	**BEEF**	**HEIF**
NTINC	1					
WOOL	.09 **	1				
LAMB	.12 **	.13 **	1			
EWE	.10 **	.39 **	.36 **	1		
BEEF	.07 *	-.12 **	.31 **	.13 **	1	
HEIF	.03	-.17 **	.33 **	.12 **	.60 **	1
TTEXP	-.32 **	.36 **	.12 **	.25 **	.03	-.02
EQ %	.26 **	.06	.17 **	.11 **	.10 **	.08 *
FT/HA	.02	-.15 **	.21 **	.02	.25 **	.20 **

SU/HA	.06 **	-.43 **	.03	-.21 **	.16 **	.19 **
LB %	.02	.14 **	-.07 *	-.04	-.07 *	-.08 **
WL/HA	.06	-.34 **	.05	-.11 **	.18 **	.27 **
AREA	-.02	.49 **	.05	.28 **	-.12 **	-.19 **
LOC.	-.05	.29 **	.004	.10 **	-.04	-.01
CLASS	-.01	-.34 **	.09 **	-.16 **	.18 **	.25 **

	TTEXP	EQ %	FT/HA	SU/HA	LB %	WL/HA
NTINC						
WOOL						
LAMB						
EWE						
BEEF						
HEIF						
TTEXP	1					
EQ %	-.10**	1				
FT/HA	.007	.14**	1			
SU/HA	-.39**	.07 **	.42 **	1		
LB %	.10 **	.03	-.04	-.29**	1	
WL/HA	-.26**	-.002	.36 **	.77 **	-.12**	1
AREA	.53 **	.05	-.21**	-.61**	.17 **	-.48**

LOC.	.27 **	-.09**	-.25**	-.59**	.26 **	-.28**
CLASS	-.28**	-.09**	.32 **	.62 **	-.20**	.58 **

	SU/HA	LB %	WL/HA	AREA	LOC.	CLASS
NTINC						
WOOL						
LAMB						
EWE						
BEEF						
HEIF						
TTEXP						
EQ %						
FT/HA						
SU/HA	1					
LB %	-.29**	1				
WL/HA	.77 **	-.12**	1			
AREA	-.61**	.17 **	-.48**	1		
LOC.	-.59**	.26 **	-.28**	.38 **	1	
CLASS	.62 **	-.20**	.58 **	-.51**	-.10**	1

* indicates significance at .01 level

** indicates significance at .001 level

Examination of the correlations between independent variables reveals interesting relationships that indicate the need for possible interaction effects within the logit model.

Expected correlations exist between WOOL, EWE and LAMB that indicates the correct sign, i.e., as the price for either increases so too does the other. Note the high correlation between the beef (steer) price (BEEF) and the heifer price (HEIF). Interaction is expected between product prices in the logit model at this stage.

Potential interaction also exists between area (AREA) and total expenses (TTEXP), stocking rate (SU/HA) and wool production (WL/HA), stocking rate and area, stocking rate and fertiliser application (FT/HA) and wool production and area.

The correlation analysis supports a qualitative assessment, or "conventional wisdoms" of some of the two-way inter-relationships found in sheep and beef farming, that may need to be included in the logit model.

Chapter Five

LOGIT MODEL SPECIFICATION

5.1 MODEL OBJECTIVES

Design of the logit model should be influenced by the proposed end use of the model. Given that the basic objective of this thesis is to develop a method for risk quantification, rather than examine or explore the 'risk' relationships between chosen variables and net cash incomes, which occurs as a consequence of this analysis anyway, then the following conditions of model building should apply.

1. The model should be as simple as possible, without detracting from the accuracy or general applicability of the final end user model. Simplicity will increase use of the model.

2. The model should be constructed such that model extension or expansion into other modes of agricultural production systems is possible without requiring alteration of any previous product type model. In other words, this model constitutes a 'micro' model for sheep and beef farming which can be incorporated into a future 'macro' model of all agricultural and aquacultural production systems. As

the first 'micro' model design must allow for extension into other forms of agricultural production systems. This would allow for the future computation of a risk index for any combination of agricultural enterprises within a production unit.

A logit modelling analysis was selected because it can operate within the thesis definition of risk and it suits the above conditions. The logit can be formulated using any dichotomous dependent variable model form. Implanting the logit into the logistic function and solving transforms the model into the 'risk index' measure by converting the expected logit, or log odds ratio, into the zero to one range, giving a direct probability measure estimate according to an assumed logistic distribution and the underlying third and fourth moments of that distribution.

To illustrate

(5.1)
$$P_i = \frac{1}{1 + e^{-L_i}}$$

is the logistic function with P_i the risk index and where $L_i = f(X_1 \ldots X_n)$ is the 'logit' and is that portion of the logistic function to be estimated before solving for P_i.

5.2 CONSTRAINTS INVOLVED IN THE MODELLING PROCESS

5.2.1 LOGIT ANALYSIS CONSTRAINTS

Although the above conditions could be considered as constraints to the modelling process, more constraining factors influence the modelling process. These constraints include the software utilised, and the availability and form of the historical data used to build the model. The analytical framework also imposes technical constraints in terms of what is statistically correct and feasible.

The design of logit analysis is governed to a large extent by sample data size. Because the process utilises a crosstabulation of transformed data from which it constructs the dependent variable and identifies the independent variables, the relationship between the size of the crosstabulation and the 'spread' of the sample data across the crosstabulation greatly influences the computability of the model. As a general rule of thumb, the average per crosstab cell size should exceed five observations, and there should be as few zero cells as is possible (Tabachnic, Fidell 1989).

Cell size is governed by not only the size of the sample data, but also by the number of variables included in the analysis and the number of qualitative classifications within each variable. The number of continuous variables, or logit covariates which are no more than normal regression variables, also influences the computability of the model through their interaction with the qualitative variables in the process.

Where continuous variables are transformed into qualitative classifications through a recoding process, then the number of classifications is governed by the division of the continuous variable into categories. The division or recoding of continuous variables is the 'weak' component of logit analysis because the divisions can be as arbitrary as one likes. In attempting to recode, which is a preliminary step in logit analysis, one wishes to apply some prior knowledge or subjective belief as to what constitutes a proper division of the continuous variable in question (McFadden 1974).

For example, consider a variable that measures the average price of an agricultural product. If the variance and range of the data is large, and as a consequence the resulting covariate coefficient derived through use of maximum likelihood within the logit modelling process is insignificant or zero, then one should convert the variable into a qualitative categorical variable if the variable must be included in the analysis.

Assume the range of the price variable in question extends from \$1.00 to \$30.00. Knowing the relationship between the number of categories within the variable and the computability of the model, how many categories should there be, and where in the continuum of the data should the divisions be made ? One could settle on three categories: good, indifferent and bad, where \$1.00 to \$10.00 may be considered 'bad', \$10.01 to \$20.00 'indifferent' and \$20.01 to \$30.00 'good'. The problem is that the entire logit process rests on such decisions, and the acceptability of the model may rely on what some may see as questionable decisions. Commentators might be inclined to question whether or not \$10.01 to \$20.00 is indeed an indifferent price range, or they might debate that three categories is insufficient to explain the 'steps' within the data range. There are no 'rules of thumb' to guide this decision process.

To reduce the arbitrary nature of this decision process, derivation, and the significance of the logit coefficients derived can be tied to this categorisation process, with the only arbitrary or subjective decision to be made being the number of qualitative categories in each independent variable.

A computer macro was developed and used to identify more objectively the appropriate ranges within each category. Using the above price variable example, the procedure was performed as follows:

1. Decide the number of categories that might best qualify the continuous variable to be transformed. In this example three categories, good, indifferent and bad, were decided upon.

2. Using a histogram of the distribution of cases in that

variable, select as a starting point three ranges that divide the sample cases reasonably evenly across the three categories. In this example \$0.00 to \$10.00, \$10.01 to \$20.00 and \$20.01 or greater, achieved an even spread. Note that the minimum \$0.00 cannot be decreased to a negative value and remains fixed because a negative gross price received is not possible as zero implies a 'non-sale'. Also note that 'greater than' implies positive infinity and therefore also remains fixed. Adjustment to the ranges within each category can only occur by altering either the \$10.00 or \$20.00 limits.

3. Systematically run logit models beginning with the ranges \$0 to \$9, \$9.01 to \$20 and \$20.01 or greater, and after each model increase the lower category limit by an increment of \$0.10. Monitor the standard error of the coefficient for the first category and select the range limit with minimum coefficient standard error. (Ideally the standard error should be smaller in absolute terms than the coefficient).

4. Once the 'correct' range limit for the first category has been identified, repeat the process keeping that first category upper limit fixed, and incrementally change the upper limit of the second category until the coefficient standard error for the second category is at a minimum. Occasionally the standard error and coefficient for the first category will change as a consequence of changing the

range for the second category. In this case select both category limits where both standard errors are jointly a minimum.

Settle at a range combination for all three categories where all standard errors for all three coefficients are, if not at a minimum, then at least smaller in size than the coefficient they relate too.

For the sake of example, assume that the final ranges for the example price variable were found to be \$0.00 to \$10.20, \$10.21 to \$21.40 and \$21.41 or greater respectively. What has essentially happened is that crosstabulation cell observations have been systematically re-distributed across the categories according to the 'best' maximum likelihood coefficient estimates for all three categories, where 'best' is defined as the coefficient with minimum standard error. This procedure is appealing because it applies elements of objectivity to the selection of range sizes. The data itself appears to indicate the 'correct' category ranges.

5.2.2 SOFTWARE CONSTRAINTS

Choice of software for the development of the logit model was governed to some extent by knowledge of software use. However, comparison of SPSSPC or SPSSX, SAS, BMDP4F, SYSTAT and SHAZAM, based on control of the modelling process and range of output reveals BMDP4F to be the most comprehensive for interaction models of association, and SPSSx to be a more general package containing both loglinear and hiloglinear components (Tabachnick, Fidel, 1989).

SPSSPC and SPSSX were utilised for the model building process, primarily because their use and output is user friendly and their 'scratch pad' command format is conducive to the command macro procedure outlined in the preceding section. SPSS does not provide an inferential test of model components but parameter estimates and their z tests are available for any specified model, along with their 95% confidence intervals. In addition, parameter estimates are reported by single degrees of freedom, so that a factor with more than two categories has no omnibus significance test reported for either its main effect or its association with other effects. Identification of an appropriate model with a large number of factors can be tedious.

Both SPSSPC and SPSSX will not accept more than ten variables, including the dependent variable. In addition, the processor memory of SPSSPC is constrained to 650K and will not allow expanded processor memory in any hardware system to be utilised. The memory requirements for the computation of any specified model is steeply exponentially related to the number of variables, and categories within variables. Memory requirements jump extraordinarily high with the inclusion of additional variables and categories. For example, up to 17mb of processing memory may be required for a simple saturated loglinear, or hiloglinear logit model consisting of ten variables, nine of which contain four categories and the dependent variable consisting of two categories (SPSSx Advanced Statistics Guide).

Specifically designed unsaturated loglinear models require relatively little processor memory. The design specifications must however incorporate all known interaction factors, previously identified using the hiloglinear system. Often a hierarchical stepping approach is required to identify the interactions prior to modelling, because of the variable constraints imposed and or the memory limitations of the hardware system used for the analysis.

5.3 MODEL DESIGN

5.3.1 CONSTRAINT CONSEQUENCES ON MODEL DESIGN

With 16 variables to analyse, it becomes obvious that with the constraints imposed by SPSSx, some will have to be eliminated early in the model building process. In order to ensure that all variables have a chance of being included within the model, the strategy employed in this thesis involved splitting the model into 'micro models', each micro model within the macro model identified or differentiated according to agricultural product, and then developing a method for combining micro models into a 'grand model'.

This strategy meets the condition of model expandability, where any future micro model can be added to a 'grand model'. Conceptually, independent linear combinations of functions summed across 'product types' to formulate the logit prior to transformation via the logistic function appealed from both the simplicity aspect as well as the expansion into other agricultural production systems.

Therefore the model consists of five 'micro models': wool, lamb, ewe, beef and heifer, where each represents the expected price variable of each product regardless of how or of what quality each is sold. For instance, wool can be sold as full fleece, second shear, fine or crossbred, differentiated only by price from a cash flow point of view. Similarly, lamb can be sold at different grades and weights, prime, store or live. Ewes, steers and heifer are also sold similarly. The model allows the farmer choice over the product mix. The model allows the farmer to choose to produce and sell only wool or lamb or ewes etc, by organising his production unit accordingly.

From the list of variables the five models to be tested initially consisted of the independent variables:

1	2	3	4	5	
WOOL	LAMB	EWE	BEEF	HEIFER	(prices)
TTEXP	TTEXP	TTEXP	TTEXP	TTEXP	
FT/HA	FT/HA	FT/HA	FT/HA	FT/HA	
SU/HA	SU/HA	SU/HA	SU/HA	SU/HA	
LB %	LB %	LB %	LB %	LB%	
WL/HA	WL/HA	WL/HA	WL/HA	WL/HA	
AREA	AREA	AREA	AREA	AREA	
LOC.	LOC.	LOC.	LOC.	LOC.	
CLASS	CLASS	CLASS	CLASS	CLASS	

with NTINC (net cash income) being the dichotomous dependent variable for each micro model. These variables were initially selected according to prior belief regarding their reflection of the components of agricultural business risk outlined in chapter one. Prices and expenses essentially reflect market risk, fertiliser use, or rather the lack of it, stocking rate, lamb sale percentage and wool quantities sold all reflect elements of production risk, and area, location and topography reflect the risk underlying physical resource constraints.

This however still leaves 10 variables per model, including NTINC, closely exceeding the software limit. Note that the above paragraph implies that the variables can be sub-divided into financial, production and physical groupings. In order to avoid the constraints of the software, a hierarchical forward selection procedure for each group was employed. That is an un-saturated logit model, with only main effects, for each group of variables was used to initially identify and eliminate 'redundant' variables. Variables were eliminated on the basis of their nil effect on the test statistic (Chi square) with their inclusion, prior to modelling fully saturated models that tested for interaction effects. Therefore no variable is included in the model unless it has a main effect on the probability estimate.

5.3.2 GENERAL MODEL SPECIFICATION

The model then is

(5.2)
$$R_i = \frac{1}{1 + e^{-L_i}}$$
where
$$L_i = \ln\left(\frac{P_i}{1 - P_i}\right)$$

Where R_i is the risk index and where the Logit L_i is the log odds ratio of $Y_i = 1$ where 1 represents zero or less net cash income. Recall from chapter three that

(5.3)
$$R_i = E(Y_i = 1 \mid X_i \ldots . X_n) = \frac{1}{(1 + e^{-L_i})}$$

where Y_i = an individual's expected net cash income given X_1 .... X_n = those independent variables upon which L_i is estimated, then a micro logit model is specified as

(5.4)
$$L_{ti} = \ln\left(\frac{P_i}{(1 - P_i)}\right) = f(\ P_{itj},\ TTEXP_{ij}, \ldots\ CLASS_{ij},\ X_{ij})$$

where P_{itj}, $TTEXP_{ij}$ ..... $CLASS_{ij}$, and X_{ij} are the independent variables, and where L_{it} = the logit for case i product price t, t = 1...n.

P_{itj} = the expected price for case i, product t, category j, j = 1...n.

$TTEXP_{ij}$ = the expected total cash expenses for case i, category j j = 1...n.

Similarly for $CLASS_{ij}$ plus other independent variables found to be significant.

X_{ij} = any interaction term for case i category j.

Note that where a variable is a covariate, j does not apply, i.e., there are no categories.

A micro logit model that contains only one categorical variable with three categories, say WOOL, (Wool price) and one continuous variable, say TTEXP, may look like

$$L_{iw} = \alpha_w + \beta_{w1} + \beta_{w2} + \beta_{w3} + \beta TTEXP_i + u_i \qquad (5.5)$$

where w1, w2, w3 are really dummy independent variables 1 or 0, and TTEXP a continuous regression variable, and u the error term.

Combination of the 'micro models' into the 'macro' sheep and beef model consists of weighting each micro model by the total income proportion of each product, and summing all weighted logits prior to solving for R_i in the logistic function. Where a farmer produces wool, lambs, ewes, and buys and sells steers, then the proportion of each in his total revenue will constitute the weights. For example, if a farmer's revenue comprises 57% wool, 22% lambs, 8% ewes and 13% steer sales then the weights on each of the five models respectively will be 0.57, 0.22, 0.08, 0.13 and 0. All weights must sum to one.

In addition, the sum of the weighted micro logits must be multiplied by two. To see this, consider a logit model with a dichotomous dependent variable. All coefficients across all categories within variables must sum to zero. For a dependent variable with two categories, say less than 0 and greater than 0, the values of the parameters are equal in absolute value but opposite in sign. Thus

$$\ln F_{11} = \alpha + \beta^1 + \beta^A + \beta^{1A} \tag{5.6}$$

which is the log of the number of those cases with zero or less than zero net cash income (= 1),

as well as

$$\ln F_{12} = \alpha + \beta^2 + \beta^A + \beta^{2A} \tag{5.7}$$

which is the log of the number of those cases with net cash incomes greater than zero (= 2), where A in both equations is some independent category of an independent variable. The log of the ratio of the two frequencies for categories 1 and 2, is the 'logit', where the log of the ratio is

$$\ln\left(\frac{F_{11}}{F_{12}}\right) = \ln F_{11} - \ln F_{12} \tag{5.8}$$

and the logit is computed for a function with only the independent variable A as

$$\begin{aligned} \ln\left(\frac{F_{11}}{F_{12}}\right) &= (\alpha - \alpha) + (\beta^{1} - \beta^{2}) \\ &\quad + (\beta^{A} - \beta^{A}) + (\beta^{1A} - \beta^{2A}) \\ &= (\beta^{1} - \beta^{2}) + (\beta^{1A} - \beta^{2A}) \end{aligned} \tag{5.9}$$

with

$$\beta^{1} = -\beta^{2} \quad \textit{as well as} \quad \beta^{1A} = -\beta^{2A} \tag{5.10}$$

Then

$$\ln\left(\frac{F_{11}}{F_{12}}\right) = 2\ (\beta^{1} + \beta^{1A})$$

Therefore the specification for the macro sheep and beef model, conditional on the variables and type of variables within each micro model, will be

(5.11) $$R_i = \frac{1}{1 + e^{-Z_i}}$$

where

(5.12) $$Z_i = 2\,[\sum_{t=1}^{n} \sigma_{it} L_{it}]$$

and where

(5.13) $$L_{it} = \alpha_{it} + \sum_{j=1}^{n} \beta_{itjv} + \sum_{j=1}^{n} \beta_{itIv} + \sum_{y=1}^{n} \beta_{ity} X_{ity}$$

conditional upon

(5.14) $$\sum_{t=1}^{n} \sigma_{it} = 1$$

and where R_i = the risk index for farmer i

Z_i = the macro logit for farmer i

σ_{it} = the proportion weights for product t

L_{it} = the micro logit for product t

α_{it} = the constant for the product t logit model

β_{itjv} = the coefficients for the j th category of variable v within the t th product model

β_{itjIv} = the coefficients for the j th category of interaction variable Iv within the t th product model

$\beta_{iy}X_{iy}$ = the coefficient times covariate y within the t th product model.

5.3.3 THE MICRO LOGIT WEIGHTS

Combination of the micro logit models is a necessary precondition in the process of developing a systematic method for expanding the macro or 'grand model'. Further examination of the addition of weighted micro logits is warranted before the individual micro models are specified. Two issues need clarification. First, is the method statistically valid, and second, why should the weights consist of product proportions?

Addressing the second question, given that net cash incomes are explained as much by total revenues as by total cash expenses, and given that within revenues both yields and price per unit sold are common stochastic independent factors across all types of agricultural production, then it makes sense to incorporate in the model a direct linkage between revenues and either the prices or yields for each product group within revenues. The linkage should consist of the proportion of total farm revenues that each product revenue contributes, where product revenue consists of the product of price times yield for that contributory product.

In addition, the product enterprises found within independent farm units, can be readily differentiated and identified according to various prices, and therefore, markets within which each farmer operates. The link therefore, between total revenue and product revenues also establishes a linkage between revenue, market price and farm yield for that product, and the combined risks inherent in both market prices and yields are incorporated in the model.

If each micro logit model represents an individual generic product group risk measure within an independent production unit, then sensibly where more than one generic product enterprise exists, the macro logit for the unit must consist of the sum of the proportional micro logits of each product the farmer is involved in.

More precisely, a logit model for lambs quantifies all the risk involved in lamb production. Should a farmer be only involved in lamb production, and nothing else, then his total farm business risk is only the business risk of lamb production. If he is also involved in wool production, and 50% of his total farm revenue is from wool sales, then his total farm business risk should be 50% of the total risk inherent in wool production and 50% of the risk inherent in lamb production. The assumption is that total risk varies according to the level of diversification found on the farm, and this method incorporates that aspect.

With regard to the first question, and referring to the above model specification, R_i, the risk index, is an estimate of $P(Y=1|Z_i)$, a conditional binomial probability. That is, R_i is an estimate of the probability that an event (Y=1) will occur conditional upon Z_i having already occurred. Since Z_i is actually the 'grand' logit, the logit can be considered an 'event' that has already occurred. P(Y=1) depends upon Z_i, the log-odds ratio, which can be a linear combination of mutually exclusive 'micro events' or micro logits. To see this, consider A (event that Y=1) and B (event Z_i). Then

(5.15)
$$P(A_i|B_i) = \frac{P(A_i \cap B_i)}{P(B_i)}$$

where

(5.16)
$$P(\bigcup_{i=1}^{n} B_i) = \sum_{i=1}^{n} P(B_i) \quad if \quad B_i \cap B_j = \emptyset$$
$$for \quad i \neq j$$

Event B can be a sum of individual events if and only if each micro event is mutually exclusive. Therefore, in this thesis it is assumed that every micro logit added to the grand logit represents an independent or mutually exclusive activity. Because a micro logit or log-odds ratio, $\ln(P_i/(1-P_i)$ is linearly related to whatever independent variable, or event, we chose to utilise, then the grand logit is merely a sum of weighted linear functions, or events, where the weights must sum to one.

Event B could consist of any number or type of mutually exclusive events so long as they constitute conditions upon which event A relies. This is the beauty of logit analysis. One could isolate just one event or condition if one so chose, and use that condition to predict the probability of an event

A occurring. Naturally the more B type events we use as conditions, the better our probability prediction of the occurrence of event A.

To see the mechanics of the proposed five micro logit models and the use of the proportion weights, expansion of the model specification formulae is necessary. Each micro logit will be numbered 1 to 5, representing wool, lamb, ewe, beef and heifer.

(5.17) $$Z_i = 2(\sum_{t=1}^{5} \sigma_{it}L_{it})$$

(5.18)
$$\frac{Z_i}{2} = \frac{\ln(\frac{P_i}{1-P_i})}{2} = \sigma_{i1}L_{i1} + \sigma_{i2}L_{i2} + \ldots + \sigma_{i5}L_{i5} + u_i$$

$$where \quad \sigma_{it} \geq 0 \qquad \sum_{t=1}^{5} \sigma_{it} = 1$$

Expanding further and dropping i for visual ease, and assuming that each micro logit function is identified by P_t the price for product t, t=1,2...5, and within each P_t model A_t and E_t variables where each variable including P_t is qualitative with two categories 1 and 2, then

$$\frac{Z}{2} = \sigma_1(\alpha_1 + \beta_{11}P_{11} + \beta_{12}P_{12} + \beta_{11}A_{11} + \beta_{12}A_{12} + \beta_{11}E_{11} + \beta_{12}E_{12})$$

$$+ \sigma_2(\alpha_2 + \beta_{21}P_{21} + \beta_{22}P_{22} + \beta_{21}A_{21} + \beta_{22}A_{22} + \beta_{21}E_{21} + \beta_{22}E_{22})$$

$$+ \sigma_3(\alpha_3 + \beta_{31}P_{31} + \beta_{32}P_{32} + \beta_{31}A_{31} + \beta_{32}A_{32} + \beta_{31}E_{31} + \beta_{32}E_{32})$$

$$+ \sigma_4(\alpha_4 + \beta_{41}P_{41} + \beta_{42}P_{42} + \beta_{41}A_{41} + \beta_{42}A_{42} + \beta_{41}E_{41} + \beta_{42}E_{42})$$

$$+ \sigma_5(\alpha_5 + \beta_{51}P_{51} + \beta_{52}P_{52} + \beta_{51}A_{51} + \beta_{52}A_{52} + \beta_{51}E_{51} + \beta_{52}E_{52})$$

(5.19)

All independent variables, being all qualitative, are either 1 or 0, yes or no dummy variables. Only one category can be selected, i.e., either P_{11} or P_{12}, A_{11} or A_{12}, E_{11} or E_{12}.

One can see that regression on the weighted variables directly would take the qualitative nature of the variables away. In logit analysis, multiplying the actual data case weights to the assigned categories would destroy the construction of the crosstabulation from which the estimated dependent odds ratio is derived.

To summarise the above expansion, we can see that

(5.20)

$$\frac{Z_i}{2} = \sum_{t=1}^{5} \sigma_t \left[\sum_{t=1}^{5} (\alpha_t + \beta_{tj}P_{tj} + \beta_{tj}A_{tj} + \beta_{tj}E_{tj}) \right]$$

where j = 1,2.

if

$$\sum_{t=1}^{5} \sigma_t = 1 \tag{5.21}$$

then

$$\frac{Z_i}{2} = \sum_{t=1}^{5} \left(\alpha_t + \beta_{tj}P_{tj} + \beta_{tj}A_{tj} + \beta_{tj}E_{tj} \right) \tag{5.22}$$

which is the basic logit functional form.

The assumption that the five micro logit models represent five mutually exclusive activities is plausible given that it is feasible to choose to conduct each activity - wool, lamb, ewe, beef and heifer - separately and exclusively from the others.

One can choose to run either sheep or cattle exclusively. If one chooses to run cattle then one can choose to either buy and sell only steers or heifers exclusively. If one chooses to run only sheep, then one can choose to buy only wethers and clip their wool and even exclude their sale if one chooses to retain them until death. One could also decide to only buy and sell ewes, not clipping their wool or mating them, or similarly only buying shorn lambs and selling them as prime full-wooled lambs. The price received for such an animal would include the value of on-the-back wool, but this is taken into account because the lamb price variable consists of a per head price received which includes slipe wool.

Accepting that the five 'sale' activities are mutually exclusive, the question remains as to whether or not the five weighted independently derived logit functions when summed, would be equal to one estimated function that included all five weighted price variables.

For example, would

$$(5.23)\qquad \sigma_1 L_{i1} = \ln\left(\frac{P_{i1}}{1 - P_{i1}}\right) = \sigma_1 (\alpha_1 + \beta_{11} X_1 + \beta_{31} X_3)$$

plus

$$(5.24)\qquad \sigma_2 L_{i2} = \ln\left(\frac{P_{i2}}{1 - P_{i2}}\right) = \sigma_2 (\alpha_2 + \beta_{12} X_2 + \beta_{32} X_3)$$

be equal to

$$(5.25)\qquad L_i = \ln\left(\frac{P_i}{1 - P_i}\right) = \alpha + \beta_1 \sigma_1 X_1 + \beta_2 \sigma_2 X_2 + \beta_3 X_3$$

where σ_1 and σ_2 are the respective proportions (weights) that X_1 and X_2 contribute to total revenue, $\Sigma\, \sigma_i = 1$, X_1 and X_2 are gross price revenue variables for products 1 and 2, and X_3 is some other variable common to both summed functions.

More specifically

$$\sigma_1(\alpha_1 + \beta_{11}X_1 + \beta_{31}X_3) + \sigma_2(\alpha_2 + \beta_{12}X_2 + \beta_{32}X_3)$$
$$= \alpha + \beta_1\sigma_1X_1 + \beta_2\sigma_2X_2 + \beta_3X_3$$

(5.26)
$$(\sigma_1\alpha_1 + \sigma_2\alpha_2) + \sigma_1\beta_{11}X_1 + \sigma_2\beta_{12}X_2 + (\sigma_1\beta_{31} + \sigma_2\beta_{32})X_3$$
$$= \alpha + \beta_1\sigma_1X_1 + \beta_2\sigma_2X_2 + \beta_3X_3$$

Therefore

$$(\sigma_1\alpha_1 + \sigma_2\alpha_2) = \alpha$$

$$\beta_{11} = \beta_1$$

(5.27)
$$\beta_{12} = \beta_2$$

$$(\sigma_1\beta_{31} + \sigma_2\beta_{32}) = \beta_3$$

An estimated equation containing all weighted variables would need to look like

$$\ln\left(\frac{P_i}{1 - P_i}\right) = (\sigma_1\alpha_1 + \sigma_2\alpha_2) + \beta_1\sigma_1X_1 + \beta_2\sigma_2X_2 + (\sigma_1\beta_{31} + \sigma_2\beta_{32})X_3$$

where the unknown parameters are unbiased maximum likelihood estimates, identical to OLS estimates, but where the MLE of the σ^2s are biased and differs from the unbiased OLS S^2s by the factor (n-k)/n (Johnston 1987).

That is, taking the natural log of the likelihood function

(5.29) $$P(y) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left[-\frac{1}{2\sigma^2} (y - \beta X)' (y - \beta X)\right]$$

then

(5.30)

$$lnL = -\frac{n}{2} \ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(y - \beta X)'(y - \beta X)$$

and differentiating partially with respect to ß and σ^2

(5.31) $$\begin{aligned} \frac{\partial(lnL)}{\partial\beta} &= -\frac{1}{2\hat{\sigma}^2}(-2X'y + 2X'X\beta) \\ &= \frac{1}{\hat{\sigma}^2}(X'y - X'X\beta) = 0 \end{aligned}$$

(5.32) $$\frac{\partial(lnL)}{\partial\sigma^2} = -\frac{1}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}(y - \beta X)'(y - \beta X) = 0$$

simultaneous solution of these k+1 equations gives

(5.33) $$\beta = (X'X)^{-1}X'y$$

which is identical to the OLS estimate, and

(5.34) $$\hat{\sigma}^2 = \frac{e'e}{n}$$

which differs from the OLS estimate

(5.35) $$S^2 = \frac{e'e}{n - k}$$

by the factor (n - k)/n.

Parameter estimates for one combined function would therefore 'decompose' to arrive at the separate micro logit model parameter estimates. Knowing the proportional weights would make this a simple exercise. Summing the qualifying coefficients from a combined function to arrive at a grand logit would be equal to the summed separate weighted micro logits.

In the case of a grand logit equation containing all variables, with each independent variable consisting of three qualitative categories, the equation would look like

(5.36)

$$L_i = (\sigma_1\alpha_1 + \sigma_2\alpha_2) + \sigma_1\begin{vmatrix}\beta_{11}X_{11}\\ \beta_{12}X_{12}\\ \beta_{13}X_{13}\end{vmatrix} + \sigma_2\begin{vmatrix}\beta_{21}X_{21}\\ \beta_{22}X_{22}\\ \beta_{23}X_{23}\end{vmatrix} + \left|\sigma_1\begin{vmatrix}\beta_{311}\\ \beta_{312}\\ \beta_{313}\end{vmatrix}\begin{matrix}X_{31}\\ X_{32}\\ X_{33}\end{matrix} + \sigma_2\begin{vmatrix}\beta_{321}\\ \beta_{322}\\ \beta_{323}\end{vmatrix}\begin{matrix}X_{31}\\ X_{32}\\ X_{33}\end{matrix}\right|$$

Here the X_{kj}s are either 1 or 0 and where L_i can qualify in only one category j or not at all.

For example a case i which does not qualify for k=2 and qualifies in category 1 (j=1) for all variables, would have logit

(5.37)
$$\begin{aligned} L_i &= \sigma_1\alpha_1 + \sigma_1\beta_{11} + \sigma_1\beta_{311} \\ &= \sigma_1(\alpha_1 + \beta_{11} + \beta_{311}) \end{aligned}$$

and a case i which qualifies for all k and category 1 (j=1) would have logit

$$(5.38)\qquad \begin{aligned} L_i &= \sigma_1\alpha_1 + \sigma_2\alpha_2 + \sigma_1\beta_{11} + \sigma_2\beta_{21} + \sigma_1\beta_{311} + \sigma_2\beta_{321} \\ &= \sigma_1(\alpha_1 + \beta_{11} + \beta_{311}) + \sigma_2(\alpha_2 + \beta_{21} + \beta_{321}) \end{aligned}$$

which is the sum of the independently weighted micro logit functions.

Chapter Six

MODEL BUILDING RESULTS

6.1 RESULTS INTRODUCTION

6.1.1 TEST STATISTICS

Test statistics reported for each model include the sample size used in the construction of each model, the MLE convergence criteria, which was set at the SPSSPC default (eps) 0.001 unless otherwise noted, the number of iterations for convergence, Pearson Chi square and its associated probability, (for saturated models Chi and its probability will always be 0 and 1 respectively), measures of entropy and concentration, and finally degrees of freedom. All important models with their associated coefficient standard errors, Z values and 95% confidence intervals are reproduced in Appendices IV, VI and VIII, along with saturated interaction effects in Appendices V and VII. The reporting sequence of the models follows the model building sequence employed in identifying the 'best' model. Crosstabulations are only reported for the final model.

Measures of entropy and concentration, or analysis of dispersion, indicate the spread about the dependent variable. Shannon's entropy measure is

(6.1)
$$H = - \sum p_i \log p_i$$

and Gini's measure of concentration is

(6.2)
$$C = 1 - \sum p_i^2$$

where P_i is the observed probability of case i occurring in NTINC Y=1. Using either of these measures it is possible to subdivide the total dispersion of the dependent variable into that explained by the model and the residual, or unexplained, variance. Entropy and concentration indicate the ratio of the dispersion 'explained' by the model to the total dispersion, and can be interpreted as measures of association. Although it is tempting to interpret the magnitudes of these measures similarly to R^2 in regression, this may be misleading since the coefficients, or measures, may be small even when the variables are strongly related (Haberman 1982). They should therefore be used within a comparative framework.

Degrees of freedom are calculated as the number of non-zero fitted cells minus the number of parameters estimated, for each model. Where empty categories, or cells, are found within the crosstabulation, the degrees of freedom of the Chi-square test may not be accurate. Therefore, given the size of the crosstabulations involved in these models, there exists a very high probability of encountering a number of empty cells for each model. Degrees of freedom measures should therefore be cautiously considered.

Where a '.' occurs within the model reports, they indicate that SPSSPC has not, or cannot, calculate the relevant statistic. Often this occurs as a default for some models. All coefficients are reported however, even where coefficients are estimated to be zero.

All models were constructed utilising SPSSPC's loglinear package, where saturated models were used to identify significant variables and interactions. Where categorisation of independent variables became necessary, recodes were initially formulated according to percentile distribution of the cases, for the testing of significant variables, and when found to significant, altered according to the previously outlined (chapter five) command macro, to establish the best mix of coefficients for each category.

Coefficient signs are interpreted as either the positive or negative effect on the log-odds ratio of a unit increase in the independent variable. Coefficients can be ranked according to sign, where negative coefficients indicate association with NTINC (2) or low or no risk, i.e., an indication of 'negative risk'. Positive signs are associated with risk, i.e., the higher the positive coefficient the greater the risk, or effect on the log-odds of belonging to NTINC (Y=1).

6.1.2 THE DEPENDENT VARIABLE

The categorisation, or recoding of net cash incomes, henceforth NTINC, once established remains unchanged for the entire model building process. NTINC is coded as; 1 = <= \$0.00, and 2 = > \$0.00, 1 and 2 having been selected instead of 0 and 1 for convenience. Coefficients reported by SPSSPC are reported for low to high codes. Therefore all coefficients reported in the appendices refer to NTINC category 1, and indicate the effect of each variable category on the log-odds of NTINC (1), the probability of actual net incomes being less than or equal to zero.

All crosstabulations reported include both NTINC (Y=1) and (Y=2), where the sum of the two categories, over all independent variables is equal to the sample size used in each model. The sample data used in this research is divided into

NTINC	(1)	(2)	Total
Number	1699	882	2581
Percent	65.8%	34.2%	100%

showing that there exists a high proportion of sample cases having shown an actual cash deficit over the five year period. This could be an indication that sheep and beef farming does indeed contain a high natural risk of realising negative cash returns.

Coefficients relating to NTINC (1) must sum to zero throughout all categories within each independent variable. Coefficients relating to the effect on the log-odds ratio for NTINC (Y=2) contain the opposite sign to those relating to NTINC (Y=1). Although SPSSPC does not report the final category coefficient for any variable, its calculation is simply determined by the opposite sign of the sum of those coefficients derived for all other categories within variables. Model reports of parameter estimates will include these derived coefficients.

6.2 COVARIATE MODEL RESULTS

The first set of models developed for testing comprise those models of the form

(6.3)

$$Ln\left(\frac{P_i}{1 - P_i}\right) = \alpha + \beta_{1j}C_j + \beta_{2j}L_j + \beta_{3j}C_jL_j + \beta_4X_4 + .. + \beta_nX_n + e_i$$

where

α = the model constant

$\beta_{1j}C_j$ = coefficient CLASS category j

$\beta_{2j}L_j$ = coefficient LOC. cat. j

$\beta_{3j}C_jL_j$ = interaction coefficient CLASS*LOC.cat. j*j

β_4X_4 .. β_nX_n = the covariate continuous variables remaining

e_i = the error term.

Various combinations and covariate transformations, along with constant and no constant models were attempted. Results of two such models are reported in Appendix IV. Model two consists of no constant and the base 10 logarithmic transformation of all independent covariates. Table 6.1 reports the statistics for each model.

Table 6.1 Models 1 & 2 statistics

	No.	Conver.	Iter.	X^2	P	DF	ENT	CONC
1	2581	.0000	20	.000	1.0	1	.011	.014
2	1194	.0000	25	.000	1.0	1	.016	.020

Note model two eliminates all cases that do not contain measures on all variables. Models containing all cases, i.e., zero indicates non-qualification in terms of producing either wool, lamb, ewe, beef or heifer, show similar results. Both models are saturated as indicated by the Chi square and its probability.

All models of this type tested contained the two qualitative variables farm class (CLASS) and regional location (LOC.). To keep the number of categories small, so avoiding too many zero cells, the codings are as follows;

CLASS	1	High or hard hill country
	2	Easier hill country
	3	Flat and intensive country
LOC.	1	Western & Central North Island
	2	Eastern North Island
	3	South Island excluding Southland
	4	Southland

Interaction between class and location (CL) reflects the regional aspects of topography. For example, C_1L_3 indicates South Island high country. These codings remain throughout the model building process. The rationale for coding LOC. is based on dividing the country into the more drought prone eastern regions from western regions, of both islands. It might be possible, on that basis, to include Westland in category 4 if indeed LOC. is found to be significant, although Westland is not represented in the data.

Model one (Appendix IV) displays what appear to be ludicrous results. First, none of the covariates can be utilised in the model, eliminating all of the price variables as well as total expenses. Second, according to the results, Southland would appear to be the most 'risky' area to farm sheep and beef in. In fact the results suggest that of all possible risk variables, a Southland location has the most risk. Clearly not acceptable results. Among the interaction terms, high country located in Southland (should it exist) has the most risk and intensive flat land the least risk, a more probable result, but not of the magnitude suggested by the model.

Transformation of the covariates and, expenses as well as area expressed in thousands and hundreds respectively, plus removal of the constant term improved the results somewhat, but again other valuable variables were not shown to hold any 'predictive' power with regard to estimating the log-odds of negative cash returns. Note the very high coefficient displayed for log base 10 WOOL, indicating that as LWOOL increases, the log-odds ratio, and therefore risk increases. Clearly not a consistent result. Similarly high or hard hill country is shown to be least risky after the interaction category hard hill or high country in Southland. Easier hill country located in Southland would seem to be the riskiest category to belong in according to model two.

Other combinations and specifications were tried in an attempt to derive a more 'correct' model. Different transformations across different covariates provided equally inaccurate results. Based on the initial results of this set of models, reverting to the five micro model design outlined in the previous chapter, using the same specification as above, resulted in equally poor results.

6.3 QUALITATIVE PRODUCTION VARIABLES MODEL

Examination of all variables to be tested reveals that the variables can be divided into three types: those that are of dollar values, i.e., prices and expenses, those that are of a physical nature, i.e., location, topographical class and farm area, and those that are descriptions of farm production, i.e., wool produced per hectare, lambing percentage, stocking rate, and to some extent fertiliser application rates.

This section deals with those variables that can be described as production variables. Their testing for significance in determining the log-odds ratio takes place within the five micro model structure, where WOOL is selected as an initial 'test micro model' incorporating all production variables and their interactions along with the average wool price variable, total cash expenses and area as continuous covariates, and the two qualitative variables, already coded, CLASS and LOC.

Models of this type will take the form

$$\begin{aligned} \ln\left(\frac{P_i}{1-P_i}\right) &= \alpha + \beta_{1j}L_j + \beta_{2j}C_j +4\ \beta_{3j}W/H_j \\ &\quad + \beta_{4j}L\%_j + \beta_{5j}S/H_j + \beta_{6j}F/H_j \\ &\quad + \beta_{7ij}X_{ij} + \beta_8 W + \beta_9 E \\ &\quad + \beta_{10}A + u_i \end{aligned} \tag{6.4}$$

where

α = the constant term

$\beta_{1j}L_{1j}$ = Coefficient for LOC. category j j=1...n

$\beta_{2j}C_{2j}$ = CLASS category j j=1...n

$\beta_{3j}W/H_{3j}$ = WL/HA category j j=1...n

$\beta_{4j}L\%_{4j}$ = LB % category j j=1...n

$\beta_{5j}S/H_{5j}$ = SU/HA category j j=1...n

$\beta_{6j}F/H_{6j}$ = FT/HA category j j=1...n

$\beta_{7ij}X_{ij}$ = interaction i i=1...n category j j=1...n

β_8W = WOOL

β_9E = TTEXP

$\beta_{10}A$ = AREA

u_i = error term.

CLASS and LOC. are naturally qualitative, and therefore will be tested along with those production type qualitative variables.

In transforming the production variables into qualitative variables the following codes were assigned:

NTINC 1 = <= $0 2 = > $0

CLASS 1 = High and hard hill country.

2 = Easier hill country

3 = Flat and intensive properties

LOC. 1 = Western and Northern North Island

2 = Eastern drought prone North Island

3 = Eastern drought prone South Island

4 = Western and Southern South Island

FT/HA 1 = Nil fertiliser application

2 = > 0 kg/ha up to 200 kg/ha

3 = > 200 kg/ha

where 2 is considered average application based on the sample distribution of application rates, and 3 considered above average.

SU/HA 1 = <= 9.6 su/ha

2 = > 9.6 su/ha

where 1 is considered below average and 2 above average.

LB % 1 = <= 95% lambing net of all deaths at end of trading

2 = > 95%

where 95% is considered a national average.

WL/HA 1 = <= 38 kg/ha

2 = > 38 kg/ha

where 38 kilograms per hectare is considered a national average, based on the sample distribution of wool production.

WOOL, TTEXP and AREA remain in their sample form, i.e., untransformed. Qualification of the above variables took into account the relationship between the number of variables and categories within variables to be tested, and the processing memory limitations imposed by SPSSPC. Inclusion of one more category within any of the above variables immediately exceeded those memory constraints. Generally the production variables are qualified according to either 'good' or 'bad' production characteristics, with the exception of FT/HA, where 'nil', 'up to average' or 'above average' seems a reasonable division.

Initial screening for all possible interaction effects between all qualitative variables was undertaken using SPSSPC Hiloglinear. Appendix V reports those findings. Based on those findings, design of the logit models included only those main and interaction effects shown to be significant.

Examination of all interactions reveals that no interaction above three way interactions are significant, with the exception of NTINC*CLASS*FT/HA*SU/HA which was ignored because it included the dependent variable NTINC. All of the significant interactions are according to expectations and are explainable. The interesting points to note are found within the nonsignificant interactions; Notably LOC.*CLASS*LB %, LOC.*FT/HA*LB %, and all other three way interactions that do not contain NTINC, and more surprisingly; FT/HA*SU/HA and

SU/HA*LB %.

In addition note that WL/HA (wool production) is a nonsignificant main effect. As a consequence of this finding, WL/HA will be eliminated from the analysis as not being significantly related to the log-odds of NTINC being <= $0.00. Inclusion of WL/HA supported the interaction findings of nonsignificance, and will no longer appear in any model.

Three models (3,4 & 5) are reported in Appendix VI, and their statistics appear in table 6.2. All models are unsaturated.

Table 6.2 Models 3,4 & 5 statistics

	No.	Conver.	Iter.	X^2	P	DF	ENT.	CONC.
3	2572	.0000	27	15.35	.951	26	.036	.041
4	2580	.0000	8	4.65	.590	6	.019	.022
5	2580	.0000	8	10.46	.314	9	.020	.024

Model 3 incorporates all those main and interaction effects identified as significant by the hiloglinear analysis. Although the 'goodness of fit' statistic for model 3 is high, (.951) examination of the statistics relating to the estimated coefficients indicate that crosstabulation problems exist due to the large number of empty cells. The relationship between P and the number of cells or categories within variables, and the number of variables, is such that the greater the number of observed empty cells, the greater the number of expected zeros within cells and therefore as a consequence the higher the goodness of fit statistic, which is misleading to say the least.

Insignificant qualitative production variables are identified from model 3. They are identified as a consequence of their very high standard errors, and or, their confidence intervals incorporating the probability of the parameter in fact being zero. Where main effects are removed from the modelling process, it has been decided that in fulfilling the objective of model simplicity their associated interactions with any other variable will also be removed.

Model 4 is the result of removing those main and interaction effects identified as nonsignificant within the model 3 construct. Goodness of fit is immediately lost as a consequence, indicating that the modelling process is still far from finished. Standard errors for coefficients of model 4 are either generally lower or similar than for the same coefficients within model 3. One interaction effect (LOC. * SU/HA) is found to be nonsignificant on the basis that its 95% confidence interval for all coefficients incorporates zero, and is removed from model 5.

One can see that as those variables that identify themselves to be nonsignificant, and are removed, them those that remain come under closer scrutiny with regard to confidence intervals and standard errors. Model 5, although providing a poor fit with the observed data, nevertheless indicates that the 'correct' variables so far are remaining with the process.

Taking each variable from model 5 in turn, and noting that the constant term has also been removed, then; LOC. (location) indicates that attention needs to be paid to the coding or classification of this variable as a consequence of the problem associated with categories 2 and possibly 4. Category 2 displays a relatively high standard error, with a 95% CI incorporating zero and the signs of the coefficients indicating that the drought prone areas of the South Island are very low risk (ie. negative sign), and Southland is very high risk (positive sign).

FT/HA (fertiliser application) shows some encouraging indication that it will remain with the process, except for the fact that it too displays the incorrect signs in terms of what is expected. The indication is that nil fertiliser application is a low risk strategy with regard to NTINC (Y=1), suggesting that the association is not toward production in terms of a greater revenue earning potential as a consequence of fertiliser application, but rather that the association is toward expenses and the risks associated with adding to total farm expenses as a consequence of fertiliser purchase.

The converse however applies to SU/HA (stock carrying capacity) where the signs are indeed correct, the standard error relatively low and a 95% CI not containing zero. The positive coefficient for a low stocking rate indicates risk in terms of NTINC (Y=1) and low risk for high stocking rates.

The interaction between LOC. and FT/HA still contains problems probably because of the problems associated with LOC.. Although it indicates that nil fertiliser application in the Western and Northern regions of the North Island (1,1) is a high risk strategy, it also indicates that nil fertiliser in the drought prone areas of the South Island (3,1) is a very low risk strategy, with a coefficient of -1.90245. In addition the log-odds ratio of NTINC (Y=1) increase for average fertiliser application in the drought prone areas of the South Island (3,2) at 1.15986, indicating high risk would appear to be related to the association of FT/HA with total farm expenses. The next stage in the process will sort this anomaly out.

The covariates WOOL (average wool price), TTEXP (total farm cash expenses) and AREA (effective farm size) show coefficient estimates that indicate problems. First WOOL shows the incorrect sign, indicating that with every dollar increase in the average wool price received the log-odds of NTINC (Y=1) increases by 4.18089 which is clearly incorrect in terms of what one would expect. The coefficient for TTEXP is very low but with a correct sign, so too is the coefficient for AREA. This might indicate that the scale of the data for both TTEXP and AREA needs to be altered. This was attempted by changing TTEXP to thousands of dollars and AREA to hundreds of hectares without much improvement to the size of the coefficients.

Overall model 5 still has a very poor fit, (.314) indicating a poor predictive power and still problems with the high residuals and number of zero cells within the crosstabulation. Variables are still not correctly identified, nor are coding ranges within variables.

6.4 ALL VARIABLES QUALITATIVE

Models of this type take the form

$$\ln\left(\frac{P_i}{1-P_i}\right) = \alpha + \beta_{1j}W_j + \beta_{2j}S_j + \beta_{3j}L_j + \beta_{4j}A_j + \beta_{5j}E_j + \beta_{6tj}X_{tj} \tag{6.5}$$

where

α = the constant term

$\beta_{1j}W_j$ = WOOL category j j = 1,2,3.

$\beta_{2j}S_j$ = SU/HA category j j = 1,2.

$\beta_{3j}L_j$ = LOC. category j j = 1,2.

$\beta_{4j}A_j$ = AREA category j j = 1,2,3.

$\beta_{5j}E_j$ = TTEXP category j j = 1,2,3.

$\beta_{6tj}X_{tj}$ = any interaction category j j = 1...n.

Transformation of the remaining continuous variables WOOL, AREA and TTEXP into categorical variables, plus a recoding of LOC. into North (1) and South (2) Islands took place prior to interaction tests. WOOL was initially coded as 1 = \$0 to \$3.00, 2 = \$3.01 to \$5.00 and 3 = greater than \$5.00, AREA code as 1 = 0 to 300 ha 2 = 300.01 to 900 ha and 3 = greater than 900 ha, and TTEXP coded as 1 = \$0.00 to \$120,000, 2 = \$120,000.01 to \$220,000 and 3 = greater than \$220,000. These recodings, although based to some extent on the distributions and statistics of each variable, are quite arbitrary. Codings will alter under closer scrutiny of the coefficients once significant and nonsignificant main and interaction effects are identified.

Interaction results are reported in Appendix VII. As expected the more obvious interactions are shown to be significant. No interaction above three way proved to be significant. Of note are the interactions between FT/HA and WOOL, FT/HA and SU/HA and FT/HA and TTEXP. Fertiliser is suspected at this stage of being more related to expenditure than revenue. Remember WOOL is not wool production, but the wool price. One could suspect that greater levels of fertiliser application could influence the quality of the wool clip so attracting a better price.

Of greater note are the nonsignificant interactions. These will be removed from the first model of this series, and consist of all three way interactions, i.e., all two way interactions need to be tested in the first loglinear unsaturated model. This model, along with an example of a subsequent model, are reported as models 6 and 7 respectively, in Appendix VIII. Table 6.3 reports the statistics of these two models.

Table 6.3 Models 6 & 7 statistics

	No.	Conver.	Iter.	X^2	P	DF	ENT.	CONC.
6	2580	.0000	22	4.715	.994	15	0.086	0.099
7	2580	.0000	18	34.652	.485	35	0.077	0.088

The initial model, model 6, shows a very good fit at 0.994. However, examination of the crosstabulation shows that a large number of zero cells is effecting this statistic, as is usually the case. Problems in the coding of nearly all the variables is also indicated by the coefficient standard errors and confidence intervals for all main effects other than SU/HA. At this stage the size of the coefficients also become concerning as they determine the positioning on the logistic cumulative function. High coefficients across all categories and variables force the resulting probability estimate range to be located at the bottom tail of the function.

Model 7 displays the results of all interaction terms deemed to be suspect and removed, with the exception of the two way interaction between AREA and TTEXP. In addition, FT/HA was proven to be associated with TTEXP in earlier model runs, and as a main effect lost significance when included in a model that contains TTEXP. It was removed from the analysis.

Model 7's poor fit again indicates both a large number of zero cells but this time combined with very high residuals, indicating that codings are incorrect. Recoding any continuous variable drastically alters the crosstabulation cell sizes. The final process is to identify what seems to be a comfortable compromise between 'fit', cell zero number, the size of residuals, the size of the coefficients and their standard errors, and a confidence interval not incorporating zero. Alteration of the codes, or inclusion/exclusion of categories and variables greatly alters the relationship between the above measures.

Of concern are the large coefficients for the constant term and WOOL. In addition, the CI for LOC. incorporates zero and its standard error is almost as big as the coefficient. LOC. is a 'pure' qualitative variable. It can not be altered through a recode. It seems tempting at this stage to remove LOC. from the analysis, under the same criteria as previous variables were removed. However, because of the nature of the variable, it is decided to persevere with it.

The confidence interval of the interaction term between TTEXP and AREA indicates a high probability of coefficients actually being zero. All CI's incorporate zero. This term will however be retained in light of the obvious relationship between farm size and fixed/variable farm expenses. Recoding of both main effects will alter the interaction coefficients. If the coefficient statistics do not change for the better, this interaction term will be removed, leaving a model with only main effects - a desirable outcome in the interests of simplicity.

6.5 THE FINAL MODEL

After identification of an acceptable WOOL micro model, all main and interaction effects for SU/HA, LOC., TTEXP and AREA will retain the same codings throughout all five micro models. Each 'price' variable will have its independent category number and coding range. By taking each WOOL model variable separately, reasonable coding ranges were identified through altering the coding ranges and re-running the loglinear program, and monitoring the relationship between all statistical output and the crosstabulation construct. Residual analysis also became an important feature in identifying a 'correct' crosstabulation.

It became apparent early in this process that alteration of the crosstab for the WOOL micro model required removal of one more variable. The sample size was not sufficiently large enough to accommodate all 'proven' variables and at the same time construct a crosstabulation that gave reasonable logit model results. Two variables came under close scrutiny. LOC. and SU/HA were obvious selections in that they both contained only two categories. LOC. codings could not be altered, but SU/HA could. Altering the codes for SU/HA indicated that it was easy to adjust the significance of the coefficients for that variable. In addition, it seemed rigorous to penalise high country South Island sheep and beef farms because of naturally much lower stocking rates in comparison to other parts of the country. SU/HA was removed as a result, and LOC. retained.

The final WOOL micro model, upon which the other four were based, takes the form

$$\ln\left(\frac{P_i}{1-P_i}\right) = \alpha + \beta_{1j}WOOL_j + \beta_{2j}LOC._j + \beta_{3j}TTEXP_j + \beta_{4j}AREA_j + u_i \tag{6.6}$$

where

α = the constant

$\beta_{1j}WOOL_j$ = Average wool price, category j j = 1,2,3.

$\beta_{2j}LOC._j$ = Island location, category j j = 1,2.

$\beta_{3j}TTEXP_j$ = Total farm cash expenses, category j j = 1,2,3.

$\beta_{4j}AREA_j$ = Effective farm area, category j j = 1,2,3.

u_i = the error term.

and where the categories j represent;

LOC.	1	=	North Island
	2	=	South Island
TTEXP	1	=	\$0.00 to \$100,000.00
	2	=	\$100,000.01 to \$150,000.00
	3	=	greater than \$150,000.00
AREA	1	=	0 ha to 400 ha
	2	=	400.001 ha to 700 ha
	3	=	greater than 700 ha.

WOOL	1	=	\$0.00 to \$3.60	per Kg.
	2	=	\$3.61 to \$4.70	
	3	=	greater than \$4.70	

for the other four micro models, $\beta_{1j}WOOL_j$ is replaced by the following;

LAMB	1	=	\$0.00 to \$15.00	per head.
	2	=	\$15.01 to \$22.00	
	3	=	greater than \$22.00	
EWE	1	=	\$0.00 to \$13.00	per head.
	2	=	greater than \$13.00	
BEEF	1	=	\$0.00 to \$450.00	per head.
	2	=	\$450.01 to \$600.00	
	3	=	greater than \$600.00	
HEIFER	1	=	\$0.00 to \$300.00	per head.
	2	=	greater than \$300.00	

The statistics for each micro model are reported in table 6.4. They indicate that all models are a 'good fit' in terms of X^2 and the number of zero cells found within each micro model crosstabulation is low. Crosstabulations are reported in Appendix IX and reasonable normality of adjusted residuals is displayed by the residual normality plots for each micro model reported in Appendix X. Tables 6.5 through 6.9 report the final set of micro models.

Cases that did not participate with regard to either wool, lamb, ewe, beef or heifer prices, i.e., they did not produce either and were deemed to be missing, were removed from the analysis in determining each micro model.

Table 6.4 Micro model statistics

	No.	Conver	Iter.	X^2	P	DF	ENT.	CONC.
WOOL	2580	.0000	5	29.73	.951	44	.074	.089
LAMB	2520	.0000	5	28.09	.962	43	.088	.106
EWE	2531	.0000	5	19.59	.879	28	.074	.090
BEEF	1731	.00002	5	37.71	.737	44	.081	.101
HEIF	1701	.0000	5	21.15	.819	28	.067	.083

All micro logit models are 'good fit' with low X^2 and significance levels high. A good fit model is one that displays no significant difference between observed cell contents and expected cell contents. Measures of association tell little of the strength of the models, except that much of the 'variance' explained is found within the residuals rather than in the models. It is difficult to know whether or not this is important, given that the models are used to predict membership to one of two categories rather than explain total variance. All models in this process displayed low entropy and concentration. But texts on SPSSPC logit analysis indicate that little attention should be given to these statistics (Tabachnic, Fidell 1989).

Table 6.5 **WOOL MICRO MODEL**

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST.		0.21031	0.02900	7.253	0.1535	0.2671
WOOL	1	0.14085	0.03847	3.661	0.0654	0.2163
	2	-0.06269	0.03128	-2.004	-0.1240	-0.0014
	3	-0.07816	.	.	.	.
AREA	1	0.42718	0.04012	10.647	0.3485	0.5058
	2	-0.14409	0.03991	-3.611	-0.2223	-0.0659
	3	-0.28309	.	.	.	.
TTEXP	1	-0.47951	0.03735	-12.838	-0.5527	-0.4063
	2	0.04569	0.03316	1.378	-0.0193	0.1107
	3	0.43381	.	.	.	.
LOC.	1	-0.08434	0.02426	-3.477	-0.1319	-0.0368
	2	0.08434	.	.	.	.

Table 6.6 **LAMB MICRO MODEL**

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST		0.14123	0.02886	4.893	0.0845	0.1978
LAMB	1	0.24781	0.03289	7.535	0.1834	0.3112
	2	-0.0547	0.03049	-1.793	-0.1144	0.0051
	3	-0.19313	.	.	.	.
AREA	1	0.48557	0.04162	11.666	0.4040	0.5672
	2	-0.15755	0.04045	-3.895	-0.2368	-0.0783
	3	-0.32803	.	.	.	.
TTEXP	1	-0.50106	0.03865	-12.964	-0.5768	-0.4253
	2	0.02595	0.03377	0.769	-0.0402	0.0921
	3	0.47511	.	.	.	.
LOC.	1	-0.10172	0.02420	-4.204	-0.1492	-0.0543
	2	0.10172	.	.	.	.

Table 6.7 **EWE MICRO MODEL**

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST		0.16188	0.02719	5.954	0.1086	0.2152
EWE	1	0.81586	0.02298	3.551	0.0366	0.1266
	2	-0.81586	.	.	.	.
AREA	1	0.43502	0.04007	10.857	0.3565	0.5136
	2	-0.14599	0.04012	-3.638	-0.2246	-0.0674
	3	-0.28903	.	.	.	.
TTEXP	1	-0.48305	0.03769	-12.817	-0.5569	-0.4092
	2	0.04028	0.03327	1.211	-0.025	0.1055
	3	0.44277	.	.	.	.
LOC.	1	-0.08558	0.02396	-3.572	-0.1325	-0.0386
	2	0.08558	.	.	.	.

Table 6.8 **BEEF (STEER) MICRO MODEL**

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST.		0.12463	0.03382	3.685	0.0583	0.1909
BEEF	1	0.14934	0.03714	4.021	0.0766	0.2221
	2	-0.07802	0.03650	-2.137	-0.1496	-0.0065
	3	-0.07802	.	.	.	.
AREA	1	0.47366	0.04960	9.550	0.3765	0.5709
	2	-0.16131	0.04517	-3.572	-0.2498	-0.0728
	3	-0.31235	.	.	.	.
TTEXP	1	-0.50543	0.04960	-10.778	-0.5974	-0.4135
	2	-0.00531	0.03895	-0.136	-0.0816	0.0710
	3	0.51074	.	.	.	.
LOC.	1	-0.09664	0.03267	-2.958	-0.1607	-0.0326
	2	0.09664	.	.	.	.

Table 6.9 **HEIFER MICRO MODEL**

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST		0.19441	0.03503	5.550	0.1258	0.2631
HEIF	1	0.08114	0.03038	2.670	0.0216	0.0141
	2	-0.08114	.	.	.	.
AREA	1	0.43163	0.04983	8.662	0.3340	0.5293
	2	-0.12302	0.04593	-2.679	-0.2130	-0.0330
	3	-0.30861	.	.	.	.
TTEXP	1	-0.46902	0.04746	-9.883	-0.5620	-0.3760
	2	-0.00153	0.03962	-0.039	-0.0792	0.0761
	3	0.47055	.	.	.	.
LOC.	1	-0.08975	0.03213	-2.794	-0.1527	-0.0268
	2	0.08975	.	.	.	.

The models generally show a middle range of 'non risk' price and expenses factors. For example, in all five models category 2, the range of expenditure from $100,000 to $150,000, indicates a relative non adjustment to the log-odds ratio. A similar null range is found within the LAMB price, where category 2 ($15.01 to $22.00) also incorporates zero within its coefficient confidence interval. In other price models, where a middle range exists, the coefficient is generally low.

Ranking of coefficients indicates that in all cases both a low effective area or a high level of farm cash expenses, on their own, are substantially higher risk factors than any other. Further, if combined then the result is one of extreme risk of incurring negative cash net returns.

Generally the five micro models fulfil the conditions set out in chapter five. Independently they are simple, meet the 'fit' criteria, have smaller coefficient standard errors in size compared to the coefficients and coefficient 95% confidence intervals generally not encompassing zero except where null ranges have been identified.

Chapter Seven

MODEL TESTING

7.1 MODEL MECHANICS

Substituting the summed proportionate logits from each micro model into the logistic equation and solving yields the risk index for the individual producer under scrutiny. Prices and all expenditures are forecasts relating to the forthcoming production year. Submitted in budget form, the details are categorised according to the ranges within each micro model. Also involved are the farm area and island location, and their coefficients belonging to the selected categories. The sum of the coefficients within each micro model are multiplied to the proportionate weights, which are obtained from either the previous years financial results or the proposed budget, summed and inserted into the logistic function.

The following is an example of how the risk index would be constructed for a hypothetical producer. Assume a farmer produces sheep and beef. The following products are sold:

Second shear crossbred wool;

The majority of his lambs as store;

The majority of his cull ewes through the saleyards;

Steers are sold prime to the works;

Cull heifers sold as 18 month stores.

Assume the hypothetical average price forecasts for the following year, i.e., the year in which seasonal finance is requested, are:

Wool	=	$ 3.28
Lamb	=	$ 13.48
Ewes	=	$ 10.20
Steers	=	$ 584.00
Heifers	=	$ 280.00

Price forecasts may be formulated using any method deemed appropriate, and may be made by either the producer or the bank or both.

Assume that this producer is located in the South Island, and the effective area upon which he produces is 485 hectares. Assume that total expenses, i.e., all cash expenses including personal expenses, or all those expenditures likely to be transacted through current account, are expected or forecast to be $ 138,794. These expenditures reflect a specific production and activity plan.

The product income proportions displayed in the producers previous accounts, and the production mix that he proposes for the following year are unchanged, and are: Wool 47%, Lambs 23%, Ewes 6%, Steers 17%, and Heifers 7%. Together these sum to 100%, regardless of how the products are sold.

Tables 7.1 and 7.2 display the calculation of the micro logits, and then the grand logit, for this example farmer. Coefficients, taken from tables 6.5 through 6.9, are selected according to the categories that the farmer qualifies in, where each category represents a range within which each of the specific price, location, area and expenses occur.

Table 7.1 Example preliminary logit calculation

		WOOL	LAMB	EWE	BEEF	HEIFER
PRICE	Val.	$3.28	$13.48	$10.20	$584	$280
	Cat.	1	1	1	2	1
	Coef.	.1409	.2478	.0816	-.0713	.0811
ISLAN	Val.	South	South	South	South	South
	Cat.	2	2	2	2	2
	Coef.	.0843	.1017	.0856	.0966	.0897
AREA	Val.	485	485	485	485	485
	Cat.	2	2	2	2	2
	Coef.	.1836	.1640	.1695	.1757	.1543

EXP.	Val.	138794	138794	138794	138794	138794
	Cat.	3	3	3	3	3
	Coef.	.0457	.0259	.0403	-.0053	-.0015
CONS.		.2103	.1412	.1619	.1246	.1944
SUM		.6648	.6807	.5389	.3203	.5181
* 2						
TOTAL		1.3297	1.3614	1.0778	0.6407	1.0362

Val. = Value   Cat. = Category   Coef. = Coefficient

Cons = Constant coefficient

Total represents the sum of the coefficients multiplied by two. At this stage the proportional weights have not been applied. Table 7.2 completes the process.

Table 7.2   Example grand logit calculation

	TOTAL	WEIGHT	MICRO LOGIT
WOOL	1.329682	0.47	0.6249505
LAMB	1.361438	0.23	0.3131307
EWE	1.077716	0.06	0.0646629
BEEF	0.640664	0.17	0.1089128
HIEFER	1.036148	0.07	0.0725303
		GRAND LOGIT	1.1841873

substituting the grand logit into the logistic function

(7.1)
$$R_i = \frac{1}{1 + e^{-L_i}}$$

where L_i is the grand logit, then

$$R_i = \frac{1}{1 + e^{-1.1841873}}$$

(7.2)
$$R_i = \frac{1}{1.305995}$$

$$R_i = 0.7656997$$

That is there is a 76.57 percent probability that this farmer will actually incur a negative net cash result at the end of his financial year, given his proposed budget.

7.2 MODEL SENSITIVITY

Completing the risk index computation for every possible category combination possible for each micro model, i.e., 51 possible combinations, where each micro model is treated as independent, i.e., all proportionate weights are one, for all micro models, results in the possible 'spread' of risk over all risk category combinations. Results of this process are reported in Appendix XI.

Table 7.3 reports the minimum and maximum risk index for each micro model.

Table 7.3 Micro model risk ranges

MICRO MODEL	P,I,A,E	MINIMUM	P,I,A,E	MAXIMUM
WOOL	3 1 3 1	0.194	1 2 1 3	0.930
LAMB	3 1 3 1	0.123	1 2 1 3	0.947
EWE	2 1 2 1	0.219	1 2 1 3	0.918
BEEF	3 1 3 1	0.149	1 2 1 3	0.938
HEIFER	1 1 3 1	0.234	1 2 1 3	0.927

P,I,A,E represent price, island, area and expenses respectively. For example, the wool micro model categories, 3 1 3 1 relates to price category 3 ($4.70 and greater), island category 1 (North Island), area category 3 (700 hectares and greater) and expenses category 1 ($0.00 to $100,000.00). All combinations are reported in the above order (PIAE).

Minimum is defined as the best possible scenario, i.e., a high price forecast, (except in the HEIF model) located in the North Island, categorisation into greater than 700 ha, and into less than $100,000 of total cash expenses, thus resulting in the lowest possible probability of sustaining negative cash net returns. Similarly maximum is defined as the worst price forecast, located in the South Island, farming less than 400 hectares, and forecast spending in excess of $150,000.

The risk ranges indicate that even when the scenario is good, there is always a probability of loss, and when bad, a probability of not sustaining loss. At 93% probability of loss, there is still a 7% probability of not sustaining loss, and equally at 20% probability of loss, there exists an 80% probability of not sustaining loss.

Appendix XI ranks each combination according to the risk index. Generally the order of risk combinations do not change greatly as risk indices proceed from lowest through highest. Of note are those combinations that display identical risk measures. Within the WOOL micro model, which assumes 100% production of wool, i.e., no other activity takes place, one sees that the PIAE categories 1232 and 3133 have identical indices. (The order 1232 is identical to the order of variables in the micro models, i.e., Price P, Location I, area A, and Expenses E. All combinations are reported in this order). Similarly 2233 and 1211, 2223 and 3112, 1212 and 3113, all have identical indices. The sensitivity risk range function for the WOOL micro model is displayed in figure 7.1. The remaining risk functions are reported in Appendix XII. Figure 7.2 indicates the range when various proportionate logit weights are incorporated. That is, 30 percent and 70 percent (0.30 and 0.70) are included to indicate that the overall slope of the step function changes according to the weight used and rotates around 0.50.

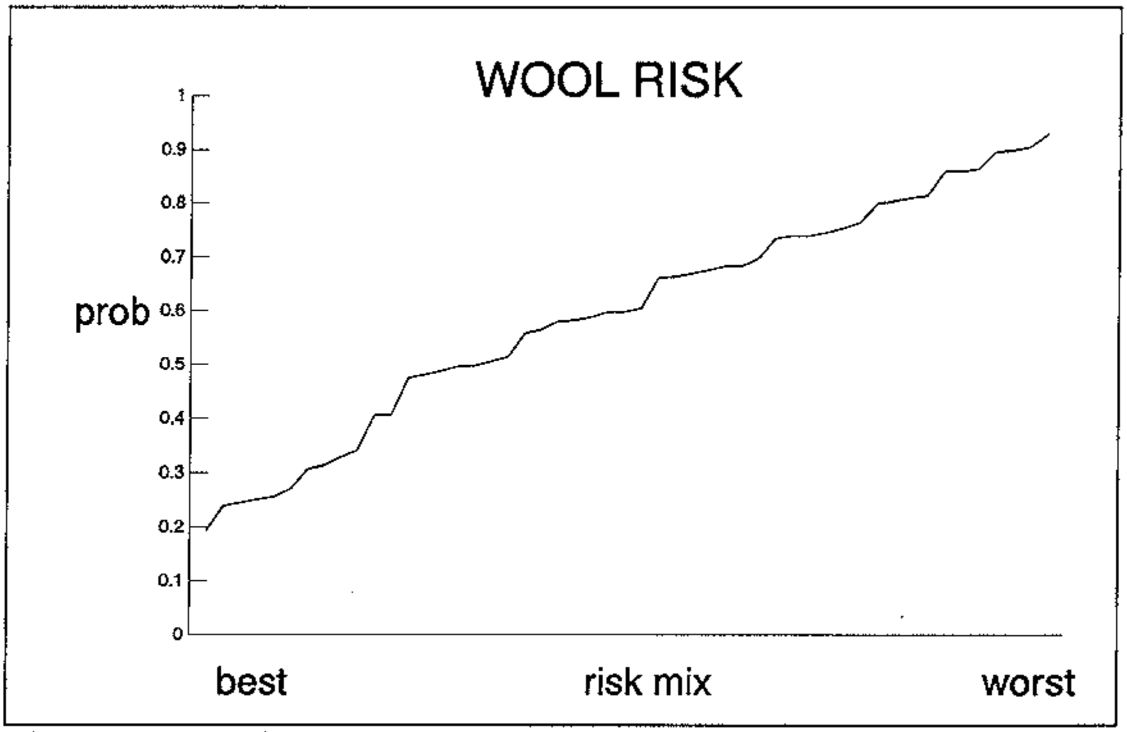


Figure 7.1 Micro wool model range

When the separate micro models are added after weighting, the risk measures can never exceed the maximum index for any individual micro model because all weights must equal one.

The resulting 'risk' function is shown to be a step function, where the steps alter slope at each significant change in categorisation. If the area to the right of the function is the area of risk, then the area to the left is the area of no risk.

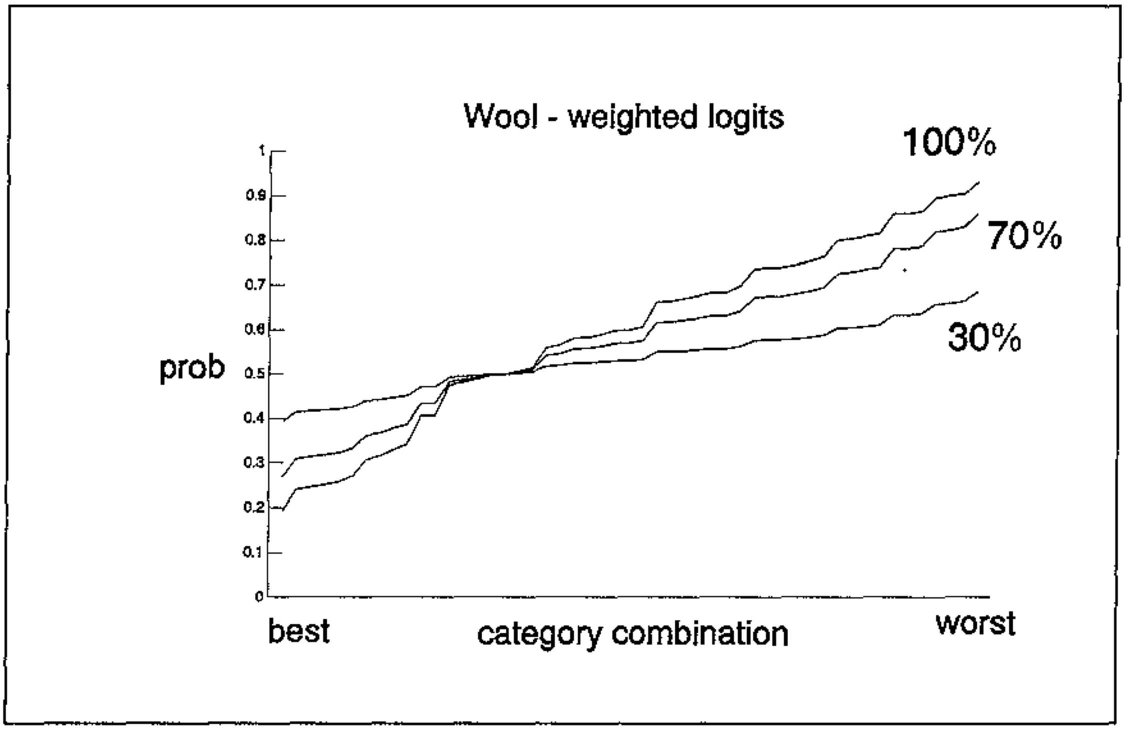


Figure 7.2 Micro wool model proportions

One can see that where a producer only produces wool, the risk of sustaining negative cash net returns is greater than the risk of not, because the majority of possible category combinations result in risk indices that exceed 0.50.

Figure 7.3 compares the risk functions for each micro model. A confusing aspect of this graph is that both the EWE and HEIF models contain only two price categories. As a consequence, the range of possible combinations is fewer than for the other three micro models. EWE and HEIF models can have only 33 category combinations.

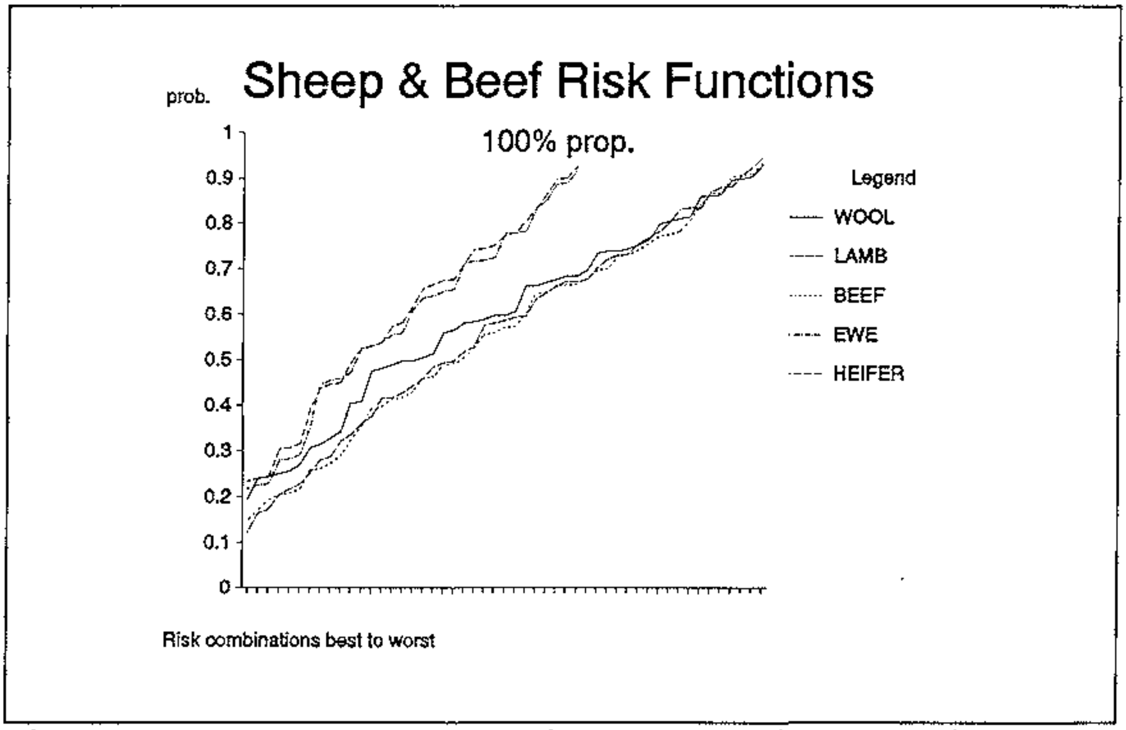


Figure 7.3 Sheep and Beef micro model risk functions

Generally the functions are nearly identical. If both EWE and HEIF are graphed on the same axis, where category price 2 is treated the same as category 3 for WOOL, LAMB and BEEF, then all functions fit along a similar area, except the steps in both EWE and HEIF are more pronounced as a consequence of having only two price categories.

The reference to risk for wool production being high is illustrated for sheep and beef farming generally when all micro models are combined. Figure 7.4 indicates the steep nature of the stepped risk function, and shows that generally the risks of sustaining negative net cash returns is high due to the high number of risk combinations with resulting risk indices exceeding 0.50. This function can therefore be considered a 'sheep and beef risk function'. The shaded area under the function indicates the region in which any individual farmer's risk index can occur according to his specific combination of categories and proportional weights.

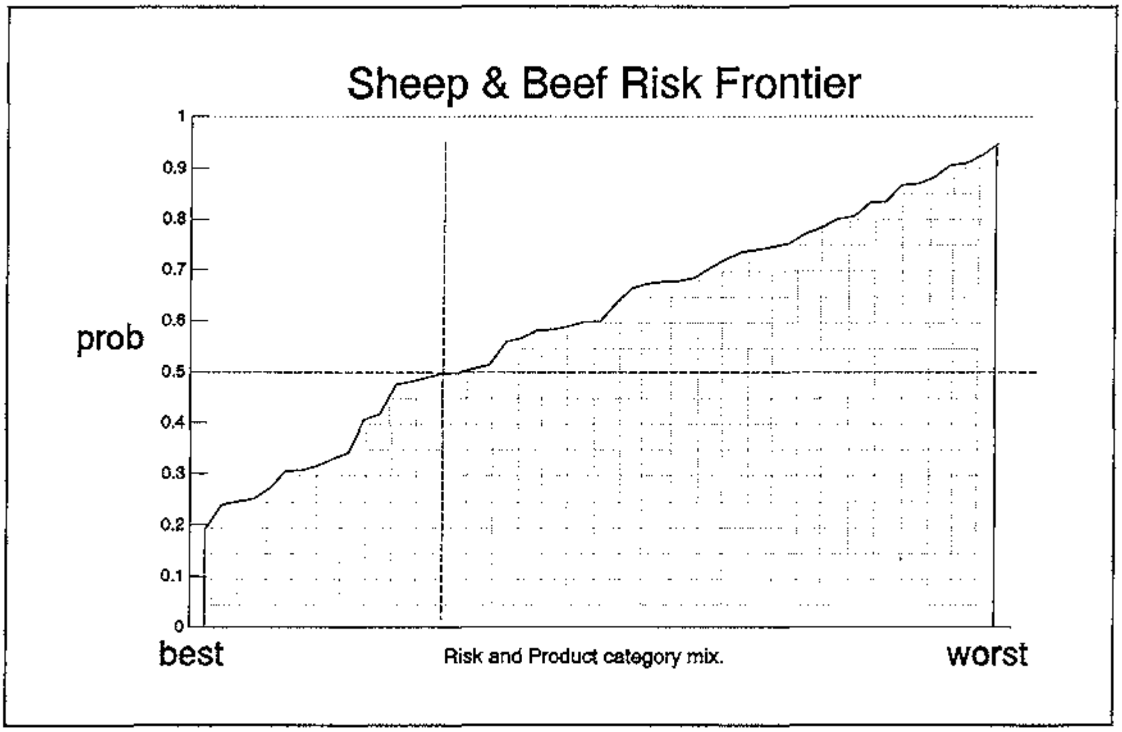


Figure 7.4 Sheep and Beef risk function

7.3 MODEL TESTING USING SAMPLE DATA

Given that the model does not predict actual cash return outcome but predicts the probability of a cash outcome, testing for 'correctness' is tantamount to testing the odds of a racehorse winning against whether or not the horse does win. If the odds of a horse winning are high, and the horse actually loses, are the odds at fault or is the horse? Because there always exists a probability of a favourable outcome, even when the probability of an unfavourable outcome is high, an actual outcome that reflects the probability of that outcome 'most of the time' could be an indication of the soundness of the probability estimate.

A random selection of 175 cases drawn from the data utilised in this study was utilised for model testing. Of the 175 cases, 113 displayed <= $0.00 net cash return, and 62 > $0.00. The risk probability measure was constructed for each of the selected cases and compared to the actual net cash return outcome of that case.

Figure 7.5 reports the distribution of the risk indices across the test sample. One notes that generally the indices are centred around the 0.48 to 0.63 area, indicating that a flip of a coin might be as good as the model. A 50 percent probability of less than zero returns could be considered no better than a flip of a coin, and measures distributed closely around 0.50 give a similar indication as to the likelihood of an outcome.

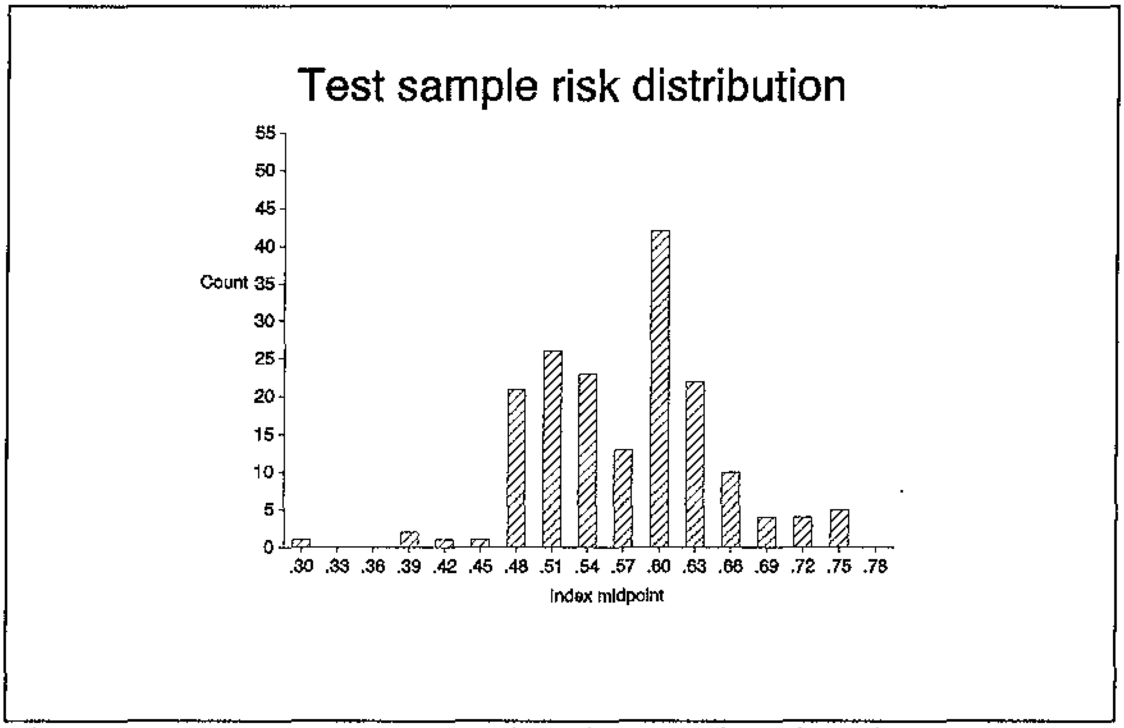


Figure 7.5 Test sample risk index distribution

Selecting the 111 cases that displayed negative net cash returns and plotting the risk distribution alongside the risk index distribution for those cases displaying positive net cash returns, as in Figure 7.6, indicate that even though the majority of risk indices give little indication as to the strength of belief one may have toward the occurrence of an outcome, the result indicates that a good proportion of the test sample that in fact had negative returns would have been predicted.

Separate risk index distributions are reported in Appendix XIII.

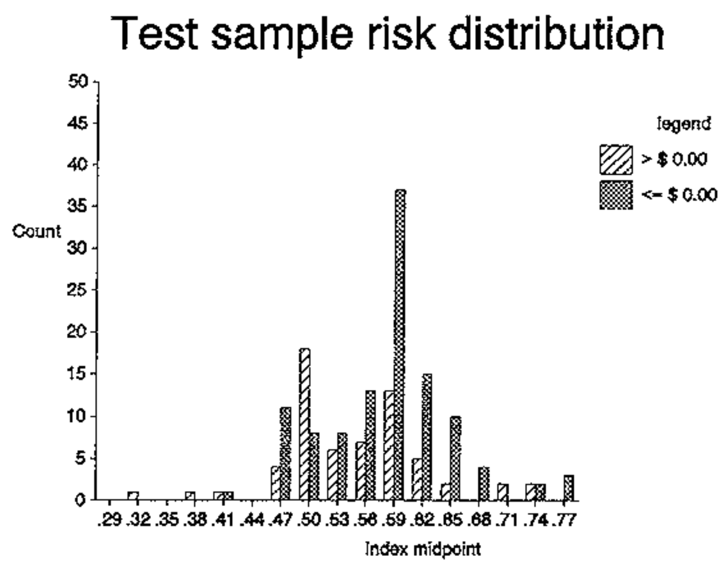


Figure 7.6 Relative risk index distribution

Those cases in the negative return category are generally distributed among the higher risk indices, as one would hope. However note the quite large overlap of both return categories across the 0.45 to 0.59 index range. This fact becomes an important feature when one is confronted with the problem of utilising the index for a yes/no decision regarding the provision of seasonal finance, a point that will be covered in a later chapter.

Assuming that a risk index greater than 50 percent would indicate a weighting or bias toward a correct probability of an unfavourable outcome, then it would be useful to see if the number of 'correct' probability indices is reflected in their ability to predict an actual unfavourable outcome. Table 7.4 indicates these results given that 50 percent is the selected threshold 'acceptance' or 'gamble' level.

Table 7.4 Probable outcome against actual

P_i <= $0	NEGATIVE	POSITIVE	TOTAL	PERCENT
> 0.40	112	60	172	65%
<= 0.40	1	2	3	66%
TOTAL	113	62	175	
> 0.50	98	46	144	68%
<= 0.50	15	16	31	52%
	113	62	175	
> 0.60	48	12	60	80%
<= 0.60	65	50	115	43%
	113	62	175	

The encouraging aspect of the above result is that the model probability estimate, where the index was greater than 0.50, selection was right 68 percent of the time, i.e., where 0.50 had been used as a reject or accept criteria, the model results would have been right to reject 68 percent of the time. However where the index is low, i.e., less than 0.50, where one would expect an outcome greater than zero net returns, the model results were right only 52 percent of the time.

If cases with a less than 0.50 index were accepted, then the model results would have accepted 31 cases, 15 of which would have turned out to be incorrect decisions (48 percent). Similarly if 144 cases were rejected the model results would have rejected 46 (28 percent) cases that would have turned out to be good clients.

Given that the intention is to use the index as an accept/reject measure, then one can obviously see that the 'correctness' of the decision rests on the threshold index one is going to use as a criteria. Should one establish 60 percent (0.60) as a threshold then fewer 'correct' acceptances would have occurred, (43 percent) but the number of 'correct' rejections would have increased, (80 percent), 20 percent of rejections constituting an opportunity loss. There appears then to be an index range within which other more substantial criteria would need to apply in the decision process involved with accepting or rejecting an application for seasonal finance. That index range appears to be around the 45 percent to 55 percent index area.

Plotting the risk indices with the actual net cash returns supports the contention that a 'null' range exists in the risk index. Figure 7.7 displays such a plot. The positioning of the scatter across the scattergram indicates the 'correctness' of a possible accept/reject decision.

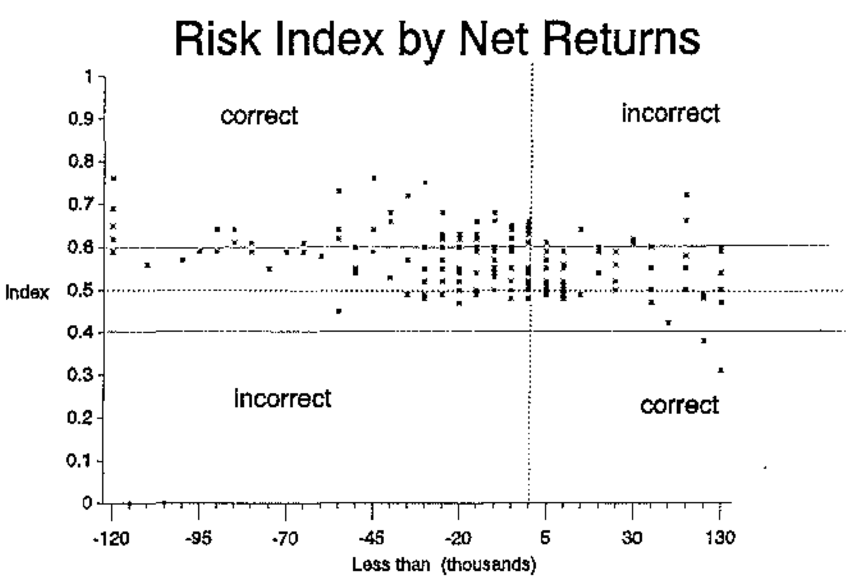


Figure 7.7 Plot of risk vs returns

One can clearly see the 'null' index range around the 50 percent level. 'Correctness' of the probability measure occurs between 0.45 and 0.55 where one would consider an 'incorrect' decision based on the index being the acceptance of a farm finance application, and in fact the outcome shown to be substantial losses.

The lower left hand quadrant indicates that no such 'mistake' occurred in the test sample. However, in the upper right hand quadrant, several 'mistakes' did occur, particularly where the index was greater than 0.6 and the final outcome indicates large cash surpluses, i.e., greater than $30,000.

These 'mistakes' would constitute an opportunity interest cost as a consequence of declination, but one could argue that an opportunity loss is better than an actual loss. Had mistakes of this magnitude been shown in the lower left hand quadrant, then the model's value would be in serious doubt.

7.4 MODEL TESTING USING ACTUAL DATA

Eight actual sheep and beef farmer bank clients were selected for risk measurement. Due to the fact that submitted budgets for the previous years trading are not usually retained, it was not possible to compare a measured risk index against an actual end of year balance. Therefore only those budgets pertaining to this year, ending June 1992, could be used. Unfortunately actual end of year balances will not be known until that time. Comparisons between the risk index and actual end of year outcomes could not therefore be made.

Nevertheless, measurements were taken on current budgets, and knowledge of opening balances used to ascertain whether or not forecast returns were in fact feasible. Budgets included price forecasts for sale produce which were made by the farmer client. In practise the risk index would utilises price forecasts that would normally be made by bank officials and not the farmer client, unless agreement existed between the bank and client as to the 'accuracy' of those forecasts. In this exercise, the index is derived using independent forecasts of price, thus eliminating a possible optimism in the farmer budget.

This exercise will indicate some of the reason and relationship between the risk index and budget forecast. Table 7.5 reports the results of risk measurement for these eight clients currently near the beginning of their production season.

Table 7.5 Actual case risk measures

CASE	F.NTINC	OPEN.BAL.	A.NTINC	AREA	F.EXP.	RISK
1	8002	-29690	21570	364	132648	65.68
2	5890	-2725	-	277	69310	51.91
3	-9276	12669	-	512	222461	60.09
4	3625	-8899	-	305	59884	52.04
5	-210	1986	-	382	78251	52.46
6	102480	-145000	-	638	276320	57.59
7	-2389	-1870	-	219	70184	51.85
8	14965	-5838	-	373	175640	72.38

Price forecast categories used in the index measure include:

Wool $ 0 to $ 3.60 = 1

Lambs $ 15.01 to $ 22.00 = 2

Ewes $ 0 to $ 13.00 = 1

Beef $ 450.01 to $ 600 = 2

Heifers $ 0 to $ 400 =1

All farms are located in the North Island.

One can see the 'null' risk index range again displayed. Those measures around the 50 percent, up to possibly 65 percent, give little strong evidence to suspect that submitted budgets are strongly at risk of being incorrect. Two cases stand alone; case 1 was one exception in not being able to obtain past data. Its measure is based on the fact that farm expenditure is high in relation to the size of the farm. The property forecast a net return (cash) of \$ 8,002 and in fact achieved a return of \$21,570. With an index measure of 65.68, and reasonable farmer forecasts for product prices, the index is explained by the 'small' size of the property, indicating that caution needs to apply in evaluating the property's ability to guarantee a positive return.

Case 8 indicates high risk. An index of 72.38, small farm size and high farm expenses, plus a lack of farm income diversification (not displayed, but 85 percent of income derived from bull beef) with extreme farmer expectations regarding the price to be received from bulls (between \$880 and \$1000 per head explaining the positive expected net income) all add up to a case that would need to undergo review of farm expenditure, as a consequence of the risk index.

Figure 7.8 plots the index against the 'forecast net cash returns'. The graph indicates those forecasts that are suspect in terms of actual outcomes. One can see that outside the null index range, two cases require examination. Cases 1 and 8, both of which have been discussed.

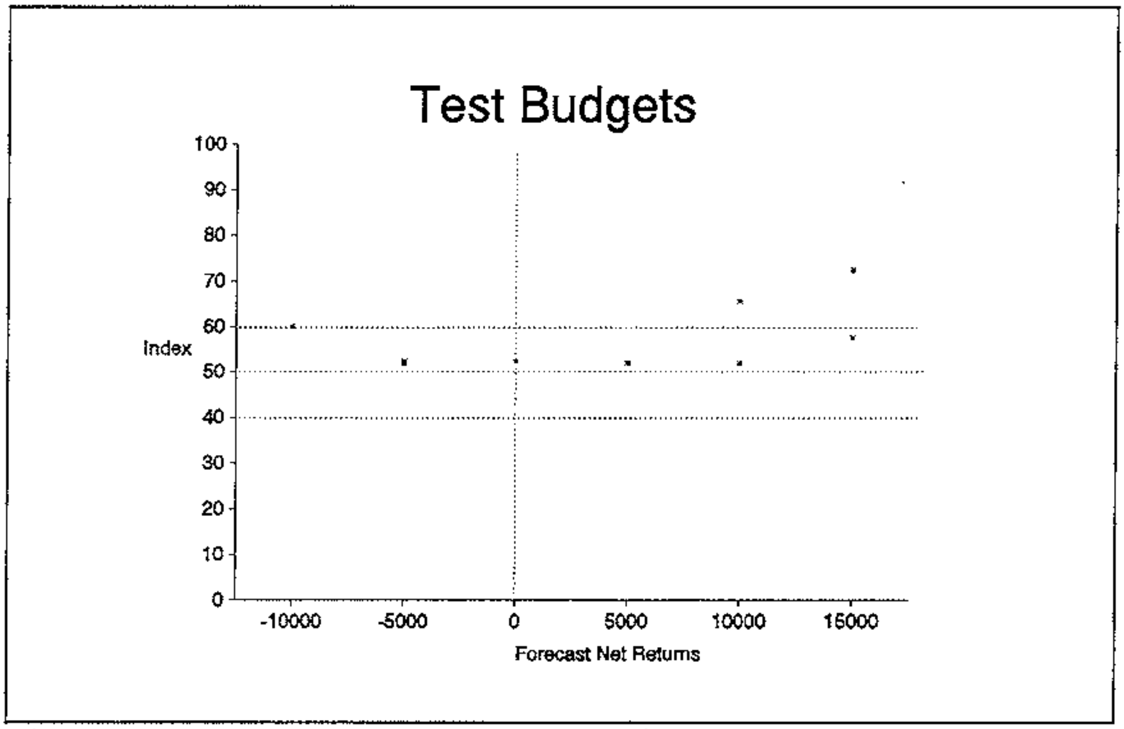


Figure 7.8 Test budgets versus risk

As a test of the 'accuracy' of the risk index, the above exercise is of no help due to the 'forecast' nature of net returns. However given the result of the test against data sample cases, one suspects that the index at the very least identifies those cases that could be potential problem clients. When used against budget forecasts of net returns, the index again indicates cases that will require a more vigorous appraisal prior to the agreement of the provision of seasonal working capital.

Chapter Eight

CONCLUSION

8.1 RESULTS DISCUSSION

The results of the modelling exercise conclusively identify and rank the most important risk variables found within sheep and beef farming. Factors such as drought, flood, policy and market instability are included implicitly within the model. These factors are displayed according to the effect that each has on those variables within the model. Drought and several peculiarities of the market are incorporated within the risk differential between North and South Islands, although the differential is not great. More important is the effect that drought, government policy and market fluctuation may have on the expenditure and price expectations of the farm plan. Drought is a factor that more often than not is included within the farms expenditure plan. The drought's effect is felt initially on farm expenditure as farmers try to protect production from the effects of the drought.

The strategy utilised in the building of the risk assessment model resulted in a straightforward model structure. Future additional micro models can be attached to the macro model.

These additional micro models need not only include main interaction effects, as did the micro models developed so far.

The resulting four variable main effect logit model was considered to be the most suitable representation of the stepped risk function for sheep and beef farming. Interaction effects were eliminated from the sheep and beef micro model only because their removal did not greatly affect goodness of fit. Goodness of fit statistics for all sheep and beef micro models were acceptable as were the confidence intervals for all coefficients.

The coefficients of the interaction term between TTEXP and AREA were small and confidence intervals indicated that for most categories zero was a highly probable coefficient. In addition, removal of SU/HA was based on the argument that codes could alter the significance of the coefficients, and penalising profitable but low stocked high country properties was not desirable.

Examination of the ranking of model coefficients reveals that within all micro models, expenditure and land size are 'greater risk' factors than average product prices. In addition, 400 hectares seems to be the threshold farm size 'step' where thereafter risks are greatly reduced. Any property engaged in sheep and beef production with an area less than 400 hectares is deemed to be high risk in the context of the model, unless the farm expenditure is less than $ 100,000, where the risk is almost totally offset. If the property is both less than 400 hectares and is prone to spending in excess of $ 150,000, then the risks are nearly doubled. If wool, lamb, ewe, beef or heifer prices are forecast to be low, then sheep and beef farmers within this farm size and expenditure category are very vulnerable to incurring negative cash returns.

This study suggests that economies of size, in terms of the relationship between expenditure and area, is the single most important risk component. The level of flexibility in being able to react to inadvertent market factors and government policy, as a result of the combination of higher than average farm expenses and limited farm size, is a key factor in establishing the 'soundness' of any application for seasonal working capital.

In addition to effective farm area, this thesis also ranks farm expenditure above average product prices in terms of impact on, or risk to net cash returns. Therefore the components of farm expenditure are more important than the components of farm revenue. Any element that greatly influences farm expenditures, whether they be physically or market sourced, seen or unforseen, must be of greater concern to the suppliers of seasonal finance than fluctuating product prices.

High price forecasts have little offsetting effect on the inherent risks underlying resource constraints such as farm size and farm expenditure limits. More often than not, favourable price forecasts will not compensate for inefficient farm expenditure, particularly within properties of less than 400 hectares, unless a major proportion of cash income is sourced outside of the sheep and beef production unit.

Large fluctuations or variation, as well as unreliability of sheep and beef product prices and costs, because of their stochastic nature, are implicitly considered the biggest risk factor within sheep and beef farming as a consequence of a general emphasis on the link between product prices, the factors that affect prices and farm profitability. The work in this thesis suggests that within reasonable price variation limits, other factors such as farm size and farm resource utilisation, have a greater bearing on the ability of a sheep and beef farmer to consequentially achieve a positive cash return.

Major farm expenditure components as well as effective farm area and the 'trade-off' between both factors are considered by this thesis to be the most important risk variables. The less is the offset to the risks inherent in small farm size by low farm expenditure, then the greater are the risks to both farmer and financial institution of farmer insolvency, keeping in mind that high product price outcomes may not be sufficient to totally offset those risks associated with joint high expenditure and small farm size.

Last but certainly by no means least, is the lowest ranking island variable differential. The island differential alters the micro logit according to what must be an inherent difference, of price, size and expenditure, and factors that affect those variables, between the two Islands. Although the 'logit' size for LOC (location) on its own is not great it has a substantial influence on the index when incorporated in the logistic function. A logit of 0.08 equates to a probability of 0.54, in comparison to -0.08 which equals 0.46, giving a difference of 0.08 or an 8 percent increase in the probability of incurring negative cash returns purely as a consequence of farming sheep and beef in the South Island.

By providing compounding levels of seasonal finance to sheep and beef farmers within the small farm size category without a low expenditure offset, regardless of which island the farmer is located, banks are in fact placing themselves at greater risk of sustaining a probable 'loan default' client. Identification of likely clients in this category is now possible with this risk index.

Constructive reaction in the form of supported diversification more suited to small farm size, or support in the purchase of additional land, given that the results of such a reaction are likely to be favourable, would seem to be a more sensible alternative to just providing ongoing seasonal overdraft facility to a property with little chance of avoiding the risks involved in producing sheep and beef on small and expensive properties.

The primary objective of this thesis was to develop a specifically agricultural risk quantification method resulting in a single index measure. This objective has been met, and a model presented indicating the process, structure, utilisation and problems of the method. Logit analysis, or the direct estimation of a risk index that defines risk as the probability of loss or harm, is a suitable replacement of the more traditional variance and monotonic transformations of variance. The index is intuitively appealing, easily understood, and apart from some probable refinement easy to construct and utilise.

8.2 MODEL STRENGTHS AND WEAKNESSES

8.2.1 LOGIT ANALYSIS TECHNIQUE

Two issues are combined in this thesis - probability as a risk measure and logit analysis as a probability estimation method. Logit analysis has been shown to be a valuable method for evaluating a 'probability' measure by directly estimating the probability of an expected outcome given independent factors. The thesis has argued that variance is not a good proxy for risk, and suggests that probability itself, because of its intuitive appeal and its implicit representation of the third and fourth moments of a distribution, is a better objective 'measure' of risk.

The main assumption that validates the method is the assumption that the cumulative distribution function, or logistic function, is representative of the CDF for the distribution that would result from the combination of multiple distributions representing the different risk factors and variables within farm cash flows. By-passing the identification of each risk factor, measuring their distributions and then deriving the moments of the combination, whether summed or multiplied, has enabled a direct estimation of implicit moments that already represent all known and unknown risk variables. One need not know how and why various factors are distributed, only that they all generally follow a similar distribution, whether normal or not.

The sample data distributions displayed in Appendix III generally indicate a skewed positive or negative log-normal distribution. Therefore use of the logistic function and the assumptions behind its use do not appear to be a gross violation. In addition, the relationships between dependent and independent variables used in the model are not linear, but generally either logarithmic or quadratic. Logit analysis assumes logarithmic relationships between variables and, apart from being convenient for estimation purposes, seems to be a reasonable method based on reasonable assumptions.

It is a conclusion of this study that, apart from some minor nuances, logit analysis is a sound technique for probability estimation of outcomes within agriculture. In addition, the ability to combine independent 'micro logits' of mutually exclusive activities into a 'macro logit' is convenient.

8.2.2 THE LOGIT MODEL

This risk index was never intended to represent an actual outcome, only the likelihood of that outcome. In the same way that variance is used to indicate the range of probable fluctuation within which an expectation will occur, that is the first and second moments of a 'net cash returns' distribution, this risk index attempts to add third and fourth moments to the information regarding an investment proposal.

Two features of logit analysis are of concern. First the relationship between both the number of categories within each variable and range of observational points within each category, and the significance or standard error of the estimated coefficient for each category within each variable. Second, the size differentials between category coefficients within each variable.

Previous discussion on the number of categories and ranges within categories indicated that apart from common sense and the sample size of data at ones disposal for such an analysis, there exist no general rules of thumb regulating the qualification of quantitative variables. The strategy attempting to identify the 'correct' range constituting each category seemed to work adequately if one accepts the criteria upon which the 'correctness' is judged. However, apart from monitoring the resulting crosstabulation once the number of categories within each variable was decided upon, and relating the number of categories to the number of zero cells within the crosstabulation as well as attempting to produce cell sizes greater than five observations, as is suggested in the literature, there is no firm rule or method for identifying the appropriate number of categories.

As can be seen from the results of the final model, the average prices variables WOOL, LAMB and BEEF have three categories, while EWE and HEIFER contain two categories. Trial and error identified the suitability of the number of categories for these variables. The result seemed to be determined by both the total range of average prices for the variables and the number of observations falling within each categorical component of the total range. A weakness then exists in that the classifications and number of categories could be deemed to be arbitrarily determined.

The size differential between category coefficients within variables was also of concern. The indication was that the 'steps' within the risk functions were large. With regard to AREA, after identification of 400 hectares as a threshold step, a farmer has a risk logit of 0.48 if he farms 398 hectares, but -0.14 if he farms 411 hectares. It seems unreasonable to suddenly penalise a farmer, or attach what amounts to be an additional 15 percent risk for the sake of one hectare about the threshold 400 hectares. Within a decision framework, two hectares could effectively lead to problems of obtaining seasonal finance if the bank adheres rigidly to an accept/decline risk threshold level.

For example, if the bank refuses to accept any application with a measured index beyond say 0.65, and a farmer with 399 hectares has an index of say 0.75 then, depending on the proportionate weights, two hectares more would decrease the index to 0.60 and the application would be accepted. This example would appear to be harsh in the extreme, and indicates that use of the index should only contribute to the information regarding the application, and should not be used as an independent criteria.

Interpolation for coefficient estimates between those estimated by logit analysis would seem to be an appropriate strategy for lessening the severity of the steps. Notice from the risk functions plotted for every possible combination, (weights = 1) that the steps in the function merely alter the slope of the function at each step. The slope is never either 1 or 0. Further, given that there can never exist a price of <= $0, then for a three category variable such as AREA or TTEXP, the first and second categories are defined points and the third category, representing a price greater than that at the second category step, defines a price that can theoretically go to infinity.

Figure 8.1 plots the coefficients for the categories within AREA of the WOOL micro model. The first category was coded lowest through 400 hectares, second 401 through 700 hectares, and third 701 through to highest area, (of the sample). Within the sample the minimum area was 53 hectares and the highest 30116 hectares. Therefore one could say that from 53 to 400 hectares the coefficient is .43, from 401 to 700 is -.15, and from 701 to 30116 ha is -.28.

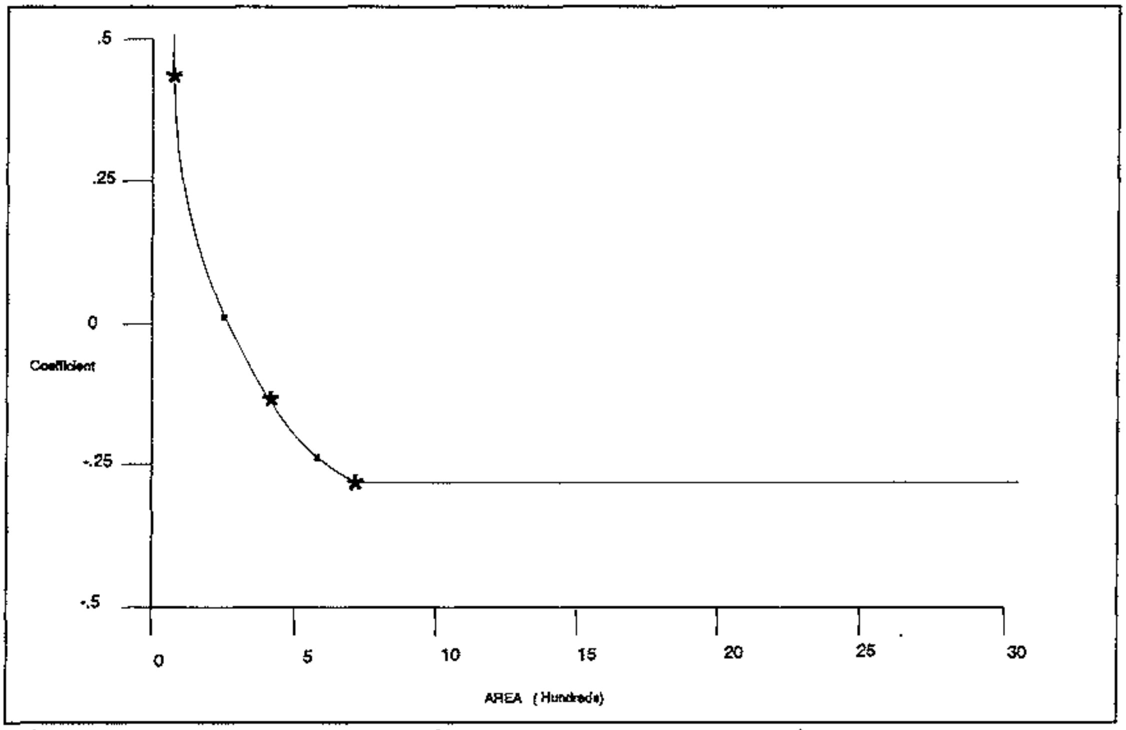


Figure 8.1 Interpolation between categories

Keeping in mind that all coefficients must sum to zero, appropriate coefficients between categories 1 and 2, and 2 and 3 could be derived mathematically using quadratic interpolation. The result would help to minimise the severity of the logit steps within the derivation of the risk index, although it is unlikely that the interpolated coefficients would be exact.

8.3 MODEL (INDEX) UTILISATION

As was previously alluded, the index is not intended to serve as a stand-alone criteria for determining the 'acceptability' of an individual farm application for seasonal finance. The information however, that is required to develop the index for an individual, must come from a budget forecast representing a specific plan of production intended for the following year, a plan which requires working capital, or a short-term seasonal overdraft facility from the bank.

The difference between cash on hand at the beginning and end of a year's trading is the best basis upon which to ascertain the likely solvency of a farmer. The index establishes a direct link and relationship between the farmer plan, his opening cash position, his proposed expenditure, and the probability of that farmer not being in a position at the end of the forthcoming years trading to service and repay the overdraft. Continued negative net cash returns, or compounding seasonal overdraft facilities, or the continued conversion of overdraft into term loan or hardcore debt, ultimately leads to a position of insolvency and bankruptcy, unless the 'probability' of this occurrence is identified at an early stage.

The index has the purpose of signalling to both farmer and bank that continued support of an ongoing long-term production plan is fraught with the risk of insolvency. A farm plan that measures high risk can then be evaluated from the perspective of both identifying those areas in the plan that are either susceptible to optimism or grossly inefficient, and or those areas in the plan that may indicate more pronounced deep seated problems regarding the resources available for such a production unit. The index may cause examination of the production unit such that conclusions regarding some form of diversification are reached, where measurement of a diversification plan reveals much less risk and greater guarantee of profits, and is therefore actively supported by the financier. The index can serve to provide confidence to both farmer and banker as to the likely outcome of any proposed venture.

An important component of the index is product price forecasts. The index does not require accurate forecasts to be applicable. It merely requires an estimate of whether or not prices are likely to be within those ranges that constitute each category. So long as one is sure that the likely price outcome will not exceed those limits imposed by the model, then one can be sure that the index will be representative of the proposed plan. Where suspicions are generated with regard to a previous forecast being inaccurate in terms of categorical classification, then a re-evaluation of the index will indicate the ability of the plan to sustain such change.

Once a proposed farm budget plan has been risk indexed, and appropriate sensitivity analysis formed on the cash flow budget to identify the maximum likely level of overdraft finance requirement, then the bank is in a position to determine, according to its portfolio objectives, the acceptability of such a proposal. Its choices are to totally decline involvement, to accept the plan unamended, to request revision of the plan, or to seek some external evaluation of the farm unit and proposed budget.

Many options are available to the bank in terms of how the index is utilised within a risk management framework. The index lends itself to utilisation within some form of Expectation-Variance (risk) decision framework, where the index replaces variance in the analysis. So long as the limits of the index are understood, i.e., that the index represents a probability estimate, and not an actual guaranteed positive or negative outcome, there is little danger that decision making will be any the worse as a consequence of it.

8.4 FURTHER RESEARCH

This research identifies two main areas where further research would be of value to this study. First, the constraints imposed by logit analysis must be lessened, thus allowing for a fine tuning of the model reported in this thesis. Second, identification of appropriate strategies, within a decision framework, for utilisation of this risk index.

The most exasperating feature of logit model building is grappling with the relationship among the size of the sample data, the number of inclusive variables and their structure, the number of categories for those quantitative variables that prove to require qualification, the range of quantitative data within each category, and the 'correctness' of the maximum likelihood coefficient estimates for every variable and category. If 'rules' can be established that ease the strategy of logit model building, then logit analysis could well become more useful than it currently is.

Throughout the model building process, the constraints imposed by logit analysis appeared to cause decisions to be made for no reason other than to make the modelling process viable. Information that has empirically proven to be 'risky', such as stocking rates and the effect of climate and resource utilisation on stocking rates, fertiliser use, the effect of international and national policy on product prices and farm expenditure, plus a host of other equally important features, were only indirectly considered in this analysis.

Use of historical data in this research assumes that all manner of factors caused fluctuation within all reported incomes and expenditures. A risk index reflecting the probability of negative cash returns is itself useful, but would be enhanced with probability estimates of the occurrence of those factors that greatly influence farm profits. A composite risk index, specifically incorporating the combined probabilities of each and every identifiable risk factor effecting farm net cash returns and solvency, could truly be called a quantification of agricultural risk.

Indices reflecting the probability of drought, increasing inflation, sudden downturns in market demand, and even the direct probability of farmer default itself, could be constructed using logit analysis. Without knowing for sure how 'correct' these probability measures are, because there seems little one can do to test the accuracy of an objective probability estimate, one is reluctant to place too much emphasis on their strength as decision criteria. Therefore further research into the logit process must incorporate sound testing methods for the identification and selection of 'correct' logit model specifications and results.

In addition to further research requirements regarding logit analysis, methods for identifying the appropriate index management techniques becomes paramount. Identification of the decision process relating to the use of the index within a banking environment is the next step in utilising a measurement of risk.

Chapter one attempted to indicate the natural bias within the relationship between bank and farmer and contended that the bank's reaction via its influence on the investment capabilities of the farmer are of at least as great a risk to the farmer than are the risks of drought or inadvertent market forces. Risk sharing strategies must first identify the risk limits that each participant is willing to accept. Because an individual farmer is more dependent on the bank's service than the bank is on an individual farmer's custom, the farmer has no choice but to operate within the risk limits imposed by the bank.

Those limits may become extended and more suitable to farmers if banks are better able to understand, measure and monitor risk, thus gaining greater confidence in agricultural investment through management of the unique risks associated with agricultural production. Quantification of agricultural risk is seen as a first step in the process of risk sharing strategies.

REFERENCES

Alderfer, LR; Bierman, H (1970)

Choices with risk: Beyond the mean and variance. Journal of Business.; 43: 341-353.

Amemiya, T (1975)

Qualitative response models. An.Ec.Soc.Measurement.; 4(3): 363-372.

Anderson, JR; Doran, HE (1978)

Some properties of simple functions of random variables. Rev.M.Ag.Ec.; 46: 37-47.

Anderson, JR; Hardaker, JB (1972)

An appreciation of decision analysis in management. Rev. of Marketing and Ag. Econ.; 40: 170-184.

Aroian, LA (1974)

The probability function of the product of two normally distributed variables. An. of Mathematics and Statistics.; 18: 265-271.

Baltensperger, E (1978)

Credit rationing: Issues and questions. Journal of Money, Credit and Banking.; 10(2): 170- 183.

Bardsley, P; Harris, M (1987)

An approach to the estimation of attitudes to risk in agriculture. Aust.J.Ag.Ec.; 31: 112-126.

Berkson, J (1951)

Why I prefer logits to probits.
Biometrics.; 7: 327-340.

Bank of New Zealand, (1990)

1990 Annual Report.

Borch, K (1967)

A note on uncertainty and indifference curves.
Rev.Ec.Studies.; 36: 1-4.

Brealey, R; Myers, S (1984)

Principles of Corporate Finance. McGraw Hill
International 2nd ed. ch.7,8,9,21,28.

Buccola, ST (1986)

Testing for nonnormality in farm net returns.
Am.J.Ag.Ec.; 68: 334-343.

Camu, L; (1977)

The daily life of a banker.
The Banker; April 1977, 67-73.

Caygill, D; Minister of Finance (1990)

Building economic growth: An economic strategy. 31-47.

Chase, SB (1960)

Credit risk and credit rationing: a comment.
Quarterly journal of Economics.; 74: 258-278.

Chew, V (1968)

Some useful alternatives to the normal distribution. Am.Statistician.; 22(3): 22-24.

Collins, RA; Barry, PJ (1986)

Risk analysis with single index portfolio models: An application to farm planning. Am.J.Ag.Ec.;68(1): 152-161.

Crosse, HD; Hempel, GH (1973)

Commercial banking functions: Management policies for commercial banks 2nd ed. Englewood CLiff,NJ. Prentice Hall. Ch.1.

Cox, D (1979)

Success in elements of banking ch.17 p.213. John Murray, London.

Day, RH (1965)

Distributions of field crop yields. Journal of Farm Economics.; 47: 713-741.

Deane, RS; Nicholl, PWE; Smith, RG (1983)

'Role of the trading bank'; Monetary policy and the New Zealand financial system. 2nd ed. Wellington NZ, Reserve bank of NZ. Ch.2.

Dillon, JL (1971)

An expository review of bernoulian decision theory in agriculture: Is utility futility? Rev.Mar.Ag.Ec.; 39: 3-80.

Elton, EJ; Gruber, MJ; Urich, TJ (1978)

Are betas best?
Journal of Finance.; 33: 1375-1384.

Feldstein, MS (1969)

Mean - variance analysis in the theory of liquidity preference and portfolio selection.
Rev.Ec.St.; 36(1): 5-11.

Francis, JC (1986)

Investments - Analysis and Management.
4th ed. McGraw Hill Book Co. Ch.9,10,12,30.

Gooding, AE (1978)

Perceived risk and capital asset pricing.
Journal of Finance.; 33: 1401-1424.

Goodman, LA (1972)

A modified multiple regression approach to the analysis of dichotomous variables.
American Sociological Review.; 37: 28-46.

Gujarati, DN (1988)

Basic Econometrics.
2nd ed. McGraw Hill International. Ch.15:467-504.

Hadar, J; Russell, WR (1969)

Rules for ordering uncertain prospects.
American Economic Review.; 59: 25-35.

Haldane, JBS (1947)

Moments of the distributions of powers and products of normal variates.
Biometrika.; 32: 209-242.

Hawke, G (1987)

'Overview of NZ Agriculture'; Rural New Zealand - What next? Ed. Wallace, LT; Latimore, R. Ag.Ec.Res.Unit; Lincoln University. Ch.1.

Hazell, PBR (1971)

A linear alternative to quadratic and semivariance programming for farm planning under uncertainty. Am.J.Ag.Ec.; 53: 53-62.

Hogdman, DR (1960)

Credit risk and credit rationing. Quarterly journal of Economics.; 74: 258-278.

Johnson, NL (1949)

Systems of frequency curves generated by methods of translation. Biometrika.; 36: 149-177.

Johnston, J (1987)

Econometric methods
3ard Edition McCraw-Hill International
Ch.7.

Just, RE (1975)

Risk response models and their use in agricultural policy evaluation. Am.J.Ag.Ec.; 57: 4836-843.

Krause, KR; Williams, PL (1971)

Personality characteristics and successful use of credit by farm families. Am.J.Ag.Ec.; 53: 619-624.

Lambert, PK; McCarl, BA (1985)

Risk modelling using direct solution of non linear approximations of the utility function.
Am.J.Ag.Ec.; 67: 846-852.

Leathers, HD; Chavas, JP (1986)

Farm debt, default and foreclosure: An economic rationale for policy action.
Am.J.Ag.Ec.; 68: 828-837.

Lee, J; Brown, DJ; Lovejoy, S (1985)

Stochastic efficiency versus mean - variance criteria as predictors of adoption of reduced tillage.
Am.J.Ag.Ec.; 67: 839-845.

Mapp, HP; Hardin, ML; Walker, OL (1979)

Analysis of risk management strategies for agricultural producers.
Am.J.Ag.Ec.; 61: 1071-1077.

McFadden, D (1974)

Conditional logit analysis of qualitative choice behaviour: Frontiers of Econometrics.
Ed. Zarembka, P. Academic Press New York.
Ch.4 p105-142.

McFadden, D (1980)

Qualitative response models: Advances in Econometrics.
Ed. Werner Hildenbrand. Papers for the 4th world congress of the Econometric Society, Aix-en-Provence.(Sept.) Ch.1:1-37.

Merrill, AS (1929)

Frequency distributions of an index when both the components follow the normal law.
Biometrika.; 20(A): 53-63.

Meyer, J (1977)

Choice among distributions.
J. Economic Theory.; 14: 326-336.

Nelson, AG; Harris, TD (1978)

Designing an instructional package: The use of probabilities in farm decision making.
Am.J.Ag.Ec.; 60: 993-997.

New Zealand Meat and Wool boards Economic Service. (1990)

NZ Sheep and Beef farm survey 1988-89.
Pub.N2020. (Nov.)

New Zealand Society of Accountants; (1985)

Management accounting for the New Zealand farmer.
NZ Society of Accountants R-404, 2nd Ed.
Wellington.

Peirson, G; Bird, R; Brown, R (1986)

Business Finance.
4th ed. McGraw Hill. Ch.6,7.

Pope, RD; (1982)

Empirical estimation and use of risk preferences:
An appraisal of estimation methods that use actual economic decisions.
Am.J.Ag.Ec.; May 1982: 376-383.

Pope, RD; Ziemer, RF (1984)

Stochastic efficiency, normality and sampling errors in agricultural risk analysis.
Am.J.Ag.Ec.; 66: 31-40.

Porter, RB; Gaumnitz, JE (1972)

Stochastic dominance versus mean - variance portfolio analysis: An empirical evaluation.
Am.Ec.Rev.; 62: 438-446.

Pryde, JG (1987)

A summary of the financial position of Canterbury farms, 1987.
A.E.R.U. Lincoln University. Dis.Pap. 114.

Pryde, JG; Bain, LB (1985)

A review of agricultural credit in New Zealand.
A.E.R.U. Lincoln University. Dis.Pap. 93.

Rae, AN; (1971)

Stochastic programming, utility, and sequential decision problems in farm management.
Am.J.Ag.Ec.; May 1971: 448-460.

Robison, T; (1987)

'Regulation'; Rural New Zealand - What next?
Ed. Wallace, LT; Latimore, R.
Ag.Ec.Res. Unit, Lincoln University; Ch.20.

Robison, LT; Barry, PJ (1977)

Portfolio adjustments: An application to rural banking.
Am.J.Ag.Ec. May 1977; 311-320.

Robison, LT; Brake, JR (1979)

Application of portfolio theory to farmer and lender behaviour.
Am.J.Ag.Ec.; 61: 158-164.

Russell, S (1985)

The open economy and competitive markets.
Reserve Bank of NZ Bulletin April 1985.

Samuelson, PA (1970)

The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments.
Rev.Ec. Studies.; 37: 537-542.

Scott, JT; Baker, CB (1972)

A practical way to select an optimum farm plan under risk.
Am.J.Ag.Ec.; 54: 657-660.

St Hill, RL (1988)

Financial market liberalisation in New Zealand: An overview.
A.E.R.U. Lincoln University. Res.Rep. 198.

Stoval, JG (1966)

Income variation and selection of enterprises.
J. Farm Econ.; 48: 1575-1579.

Tabachnic, BG; Fidell, LS (1989)

Using Multivariate statistics.
2nd ed. Harper and Row New York. Ch.7:236-315.

Taylor, CR (1986)

Risk aversion versus expected profit maximisation with a progressive income tax.
Am.J.Ag.Ec. Feb 1986; 137-143.

Tsiang, SC (1972)

The rationale of the mean - Standard deviation analysis, skewness preference and the demand for money.
Am.Ec.Rev.; 62: 354-371.

Turvey, CG; Driver, HC (1987)

Systematic and nonsystematic risks in agriculture.
Can.J.Ag.Ec.; 35: 387-401.

Van Horne, J; Nicol, R; Wright, K (1981)

Financial management and policy in Australia.
Prentice - Hall of Australia 1981; Ch.7.

APPENDIX I.

DATA VARIABLE FIELDS.

1. Farm Class
2. Survey Year
3. Wool Sales $
4. Wool Sold Kg
5. Prime Lamb Sales $
6. Prime Lambs Sold
7. Prime Mutton Sales $
8. Prime Mutton Sold
9. Prime Steer Sales $
10. Prime Steers Sold
11. Prime Heifer Sales $
12. Prime Heifers Sold
13. Store Whether/Ram/Cry. Lamb Sales $
14. Store Whether/Ram/Cry. Lambs Sold
15. Store Ewe Lamb Sales $
16. Store Ewe lambs Sold
17. Live Ram/Ewe Lamb Sales $
18. Live Ram/Ewe Lambs Sold
19. Store CFA Ewe Sales $
20. Store CFA Ewes Sold
21. Store 2T Ewe Sales $
22. Store 2T Ewes Sold
23. Store 2yr Steer Sales $
24. Store 2yr Steers Sold
25. Store 1yr Steer sales $

26. Store 1yr Steers Sold
27. Store 2yr Heifer sales $
28. Store 2yr Heifers Sold
29. Store 1yr Heifer Sales $
30. Store 1yr Heifers Sold
31. Weaner Steer sales $
32. Weaner Steers Sold
33. Weaner Heifer sales $
34. Weaner heifers Sold
35. Total Cash Income
36. Total Working Expenses
37. Capital Purchases
38. Personal Drawings
39. Interest Payments
40. Fertiliser Expenses
41. Taxation payments
42. Mortgage Payments
43. Cash Equity
44. Total Asset values
45. Phosphate Fertiliser Applied tonnes
46. Nitrogenous Fertiliser Applied tonnes
47. Effective farm Area
48. Regional location
49. Ewe Numbers
50. 2T Ewe Numbers
51. Ewe Hogget Numbers
52. Whether/Ram Hogget Numbers
53. Ram Numbers

54. Mated Cow/Heifer Numbers
55. Yearling heifer Numbers
56. Yearling steer Numbers
57. 2yr Steer Numbers
58. 3yr Steer Numbers
59. Bull Beef Numbers
60. Breeding Bull numbers

APPENDIX II.

COUNTIES AND DISTRICTS WITHIN GEOGRAPHICAL REGIONS

REGION 1. NORTHLAND

Mangonui

Whangaroa

Hokianga

Bay of Islands

Whangarei

Hobson

Otamatea

Rodney

Waiheke

REGION 2. WAIKATO /BOP /CENTRAL PLATEAU /COROMANDEL

Great Barrier Island

Franklin

Raglan

Waikato

Waipa

Otorohanga

Waitomo

Taumaranui

Thames - Coromandel

Hauraki

Ohinemuri

Piako

Matamata

Tauranga

Rotorua

Taupo

Whakatane

Opotiki

Waimarino

REGION 3. EAST CAPE

Waiapu

Waikohu

Cook

REGION 4. HAWKES BAY

Wairoa

Hawke's Bay

Waipawa

Waipukurau

Dannevirke

Woodville

Pahiatua

Eketahuna

REGION 5. WAIRARAPA

Masterton

Wairarapa

Featherston

REGION 6. TARANAKI /MANAWATU /WELLINGTON

Clifton

North Taranaki

Inglewood

Stratford

Egmont

Eltham

Waimate West

Hawera

Patea

Waitotara

Wanganui

Rangitikei

Kiwitea

Pohangina

Oroua

Manawatu

Kairanga

Horowhenua

Hutt

REGION 7. WESTCOAST

Golden Bay

Waimea

Buller

Inangahua

Grey

Westland

REGION 8. MARLBOROUGH

Marlborough

Kaikoura

REGION 9. CANTERBURY

Amuri

Cheviot

Hurunui

Rangiora

Eyre

Oxford

Malvern

Paparau

Waimairi

Heathcote

Mount Herbert

Akaroa

Chathams

Wairewa

Ellesmere

Ashburton

Strathallan

Mackenzie

Waimate

Waitaki

Waihemo

REGION 10. OTAGO

Silverpeaks

Bruce

Clutha

Tuapeka

Maniototo

Vincent

Queenstown

REGION 11. SOUTHLAND

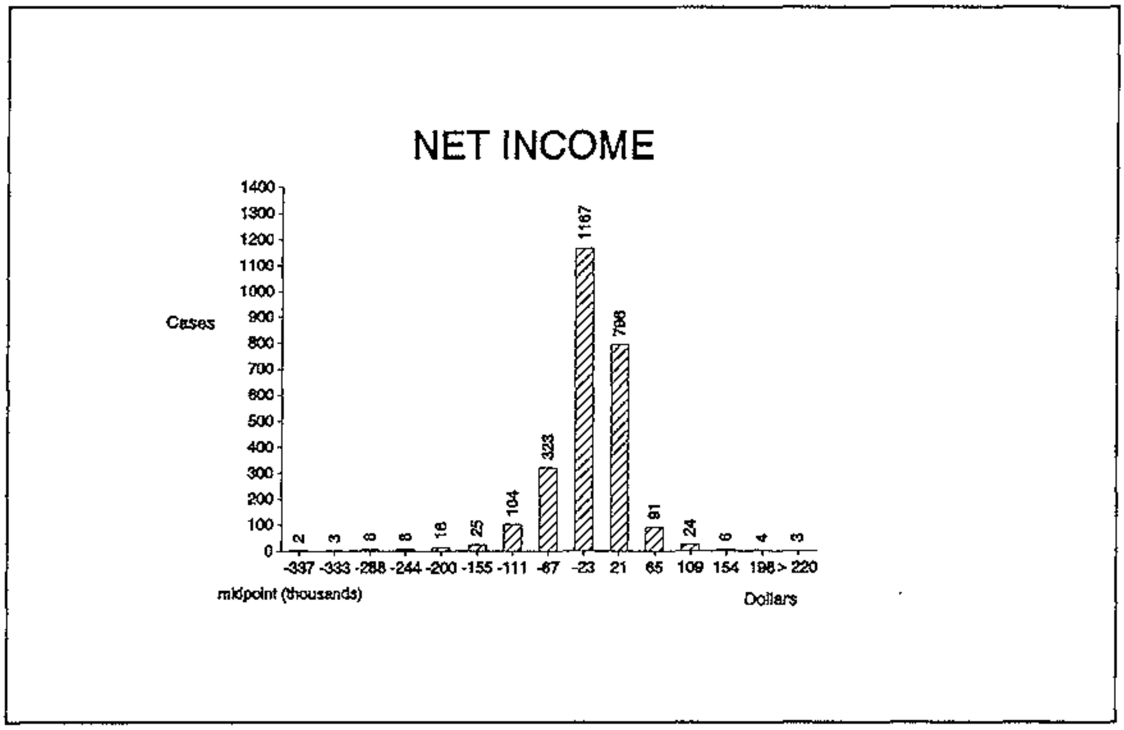
Southland

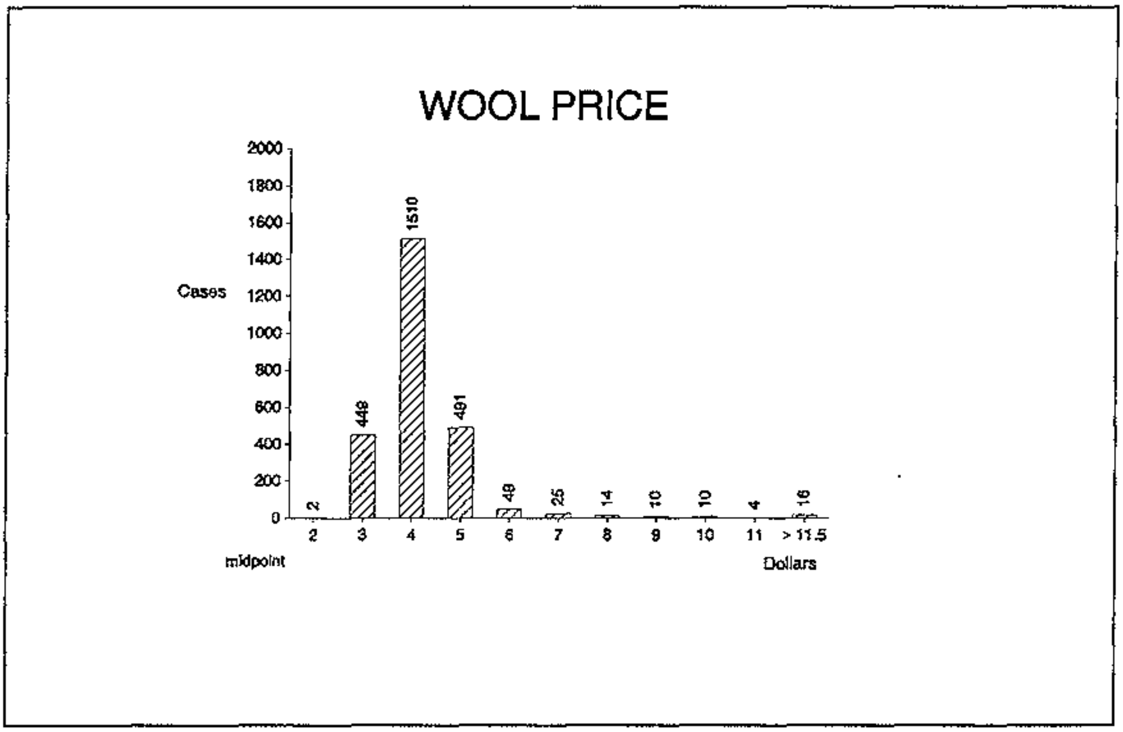
Wallace

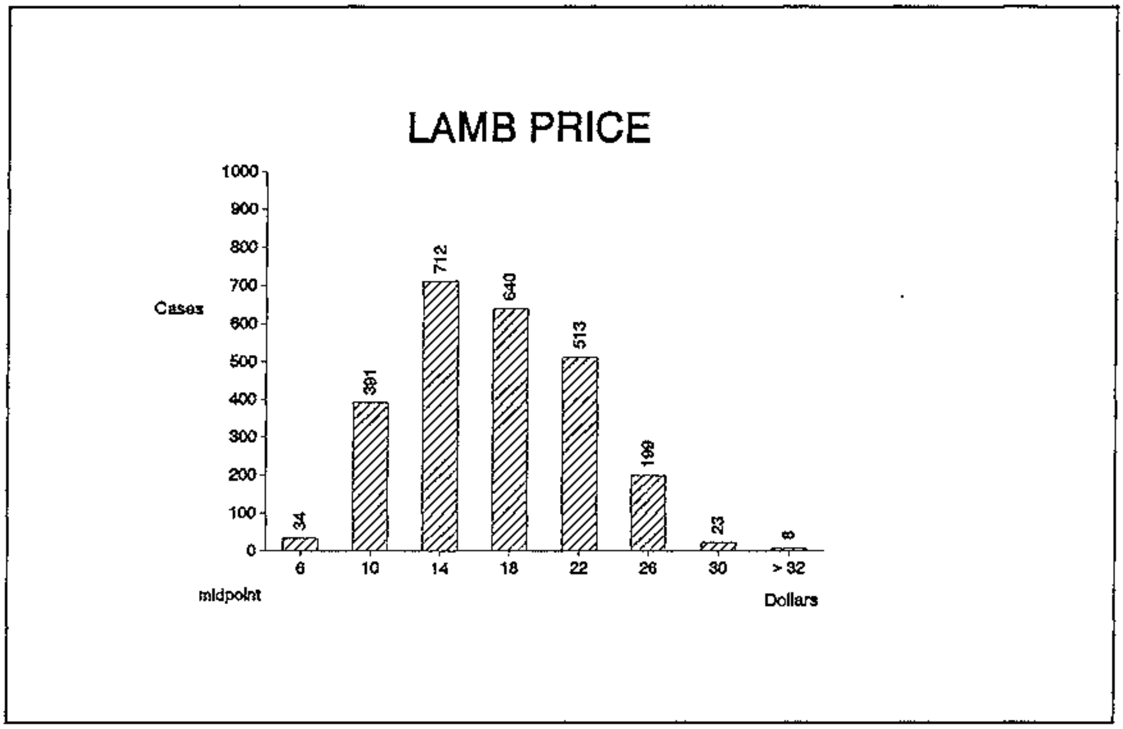
Stewart Island

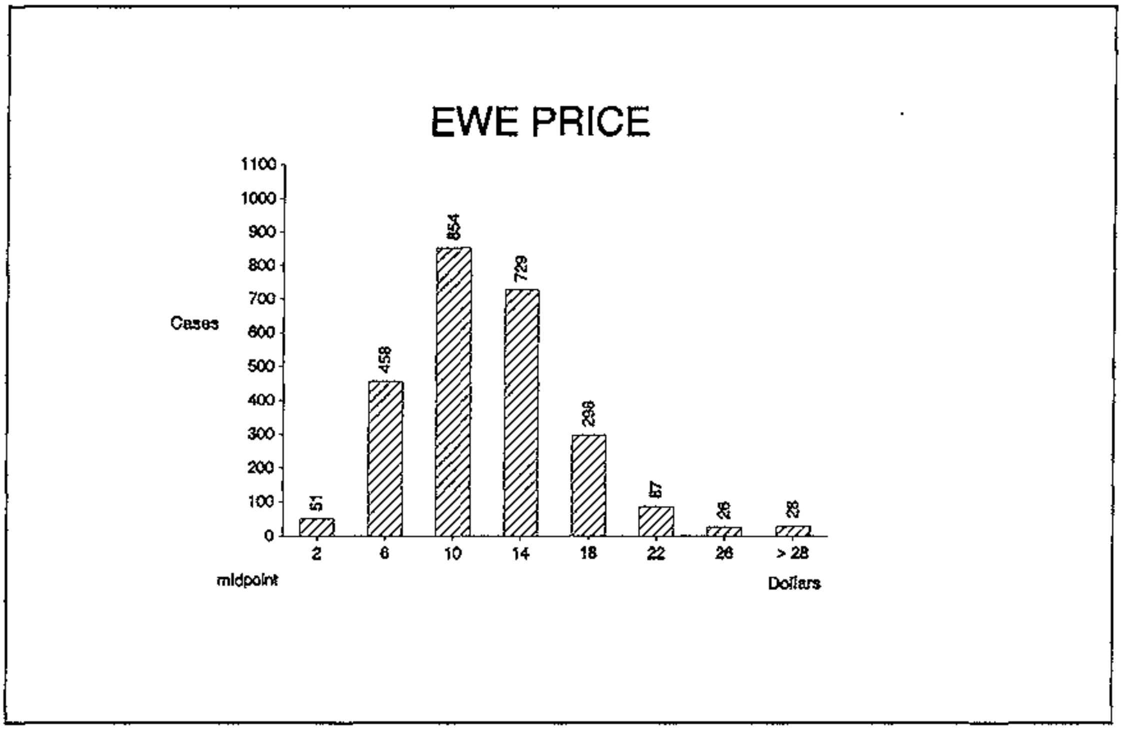
APPENDIX III.

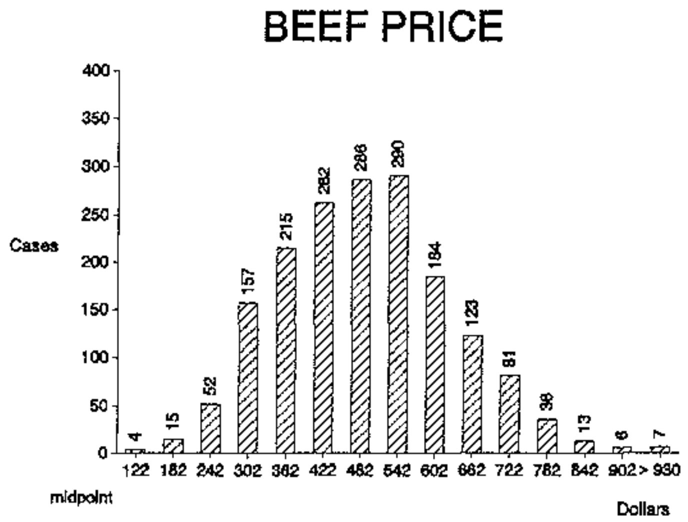
VARIABLE DISTRIBUTIONS

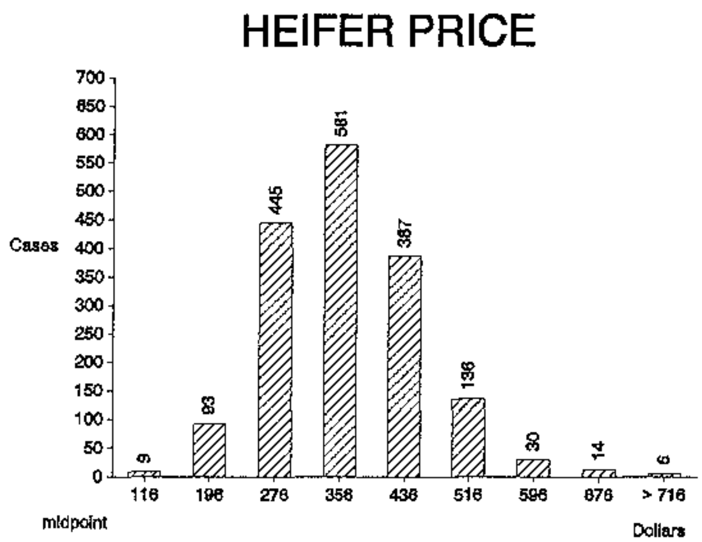


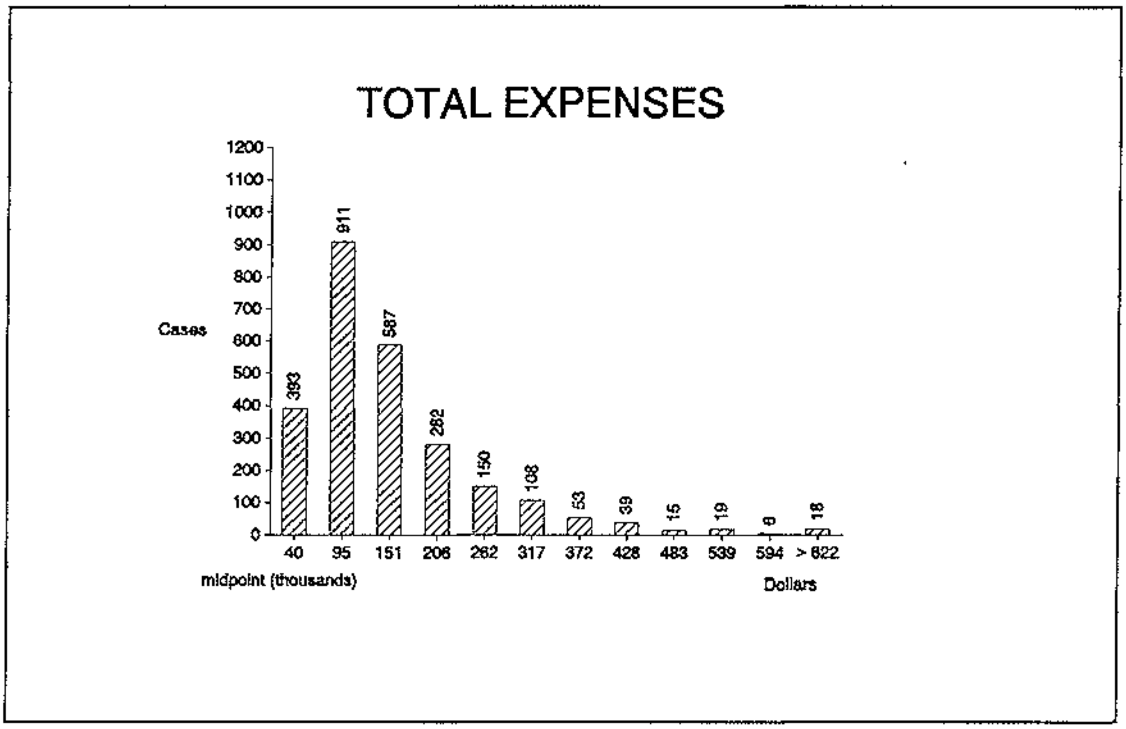


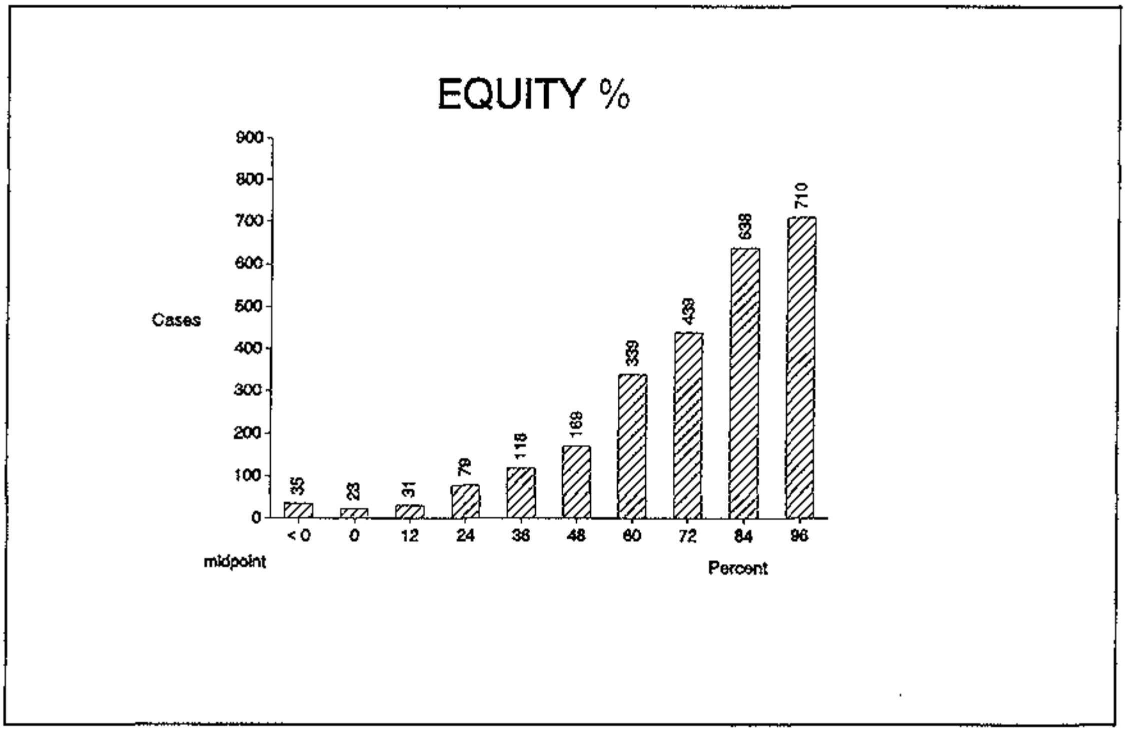

LAMB PRICE
Cases
1000
900
800
700
600
500
400
300
200
100
0
34
391
712
640
513
199
23
8
6
10
14
18
22
26
30
> 32
midpoint
Dollars

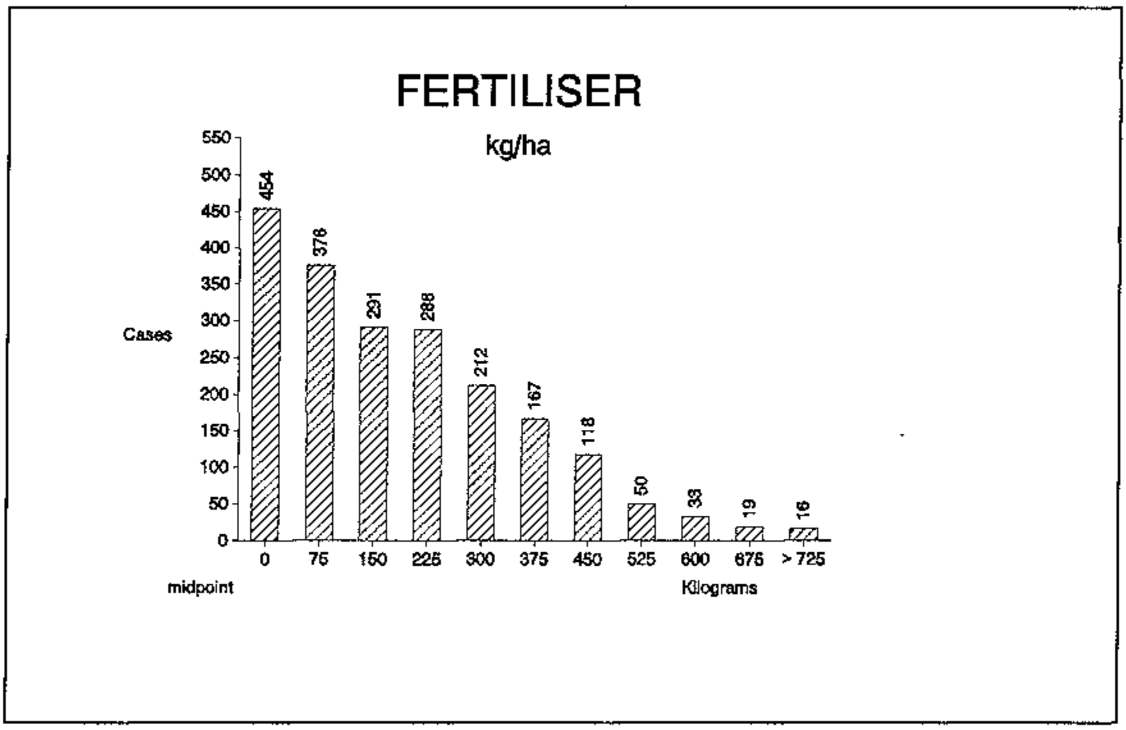

EWE PRICE
Cases
1100
1000
900
800
700
600
500
400
300
200
100
0
51
458
854
729
298
87
26
28
2
6
10
14
18
22
26
> 28
midpoint
Dollars

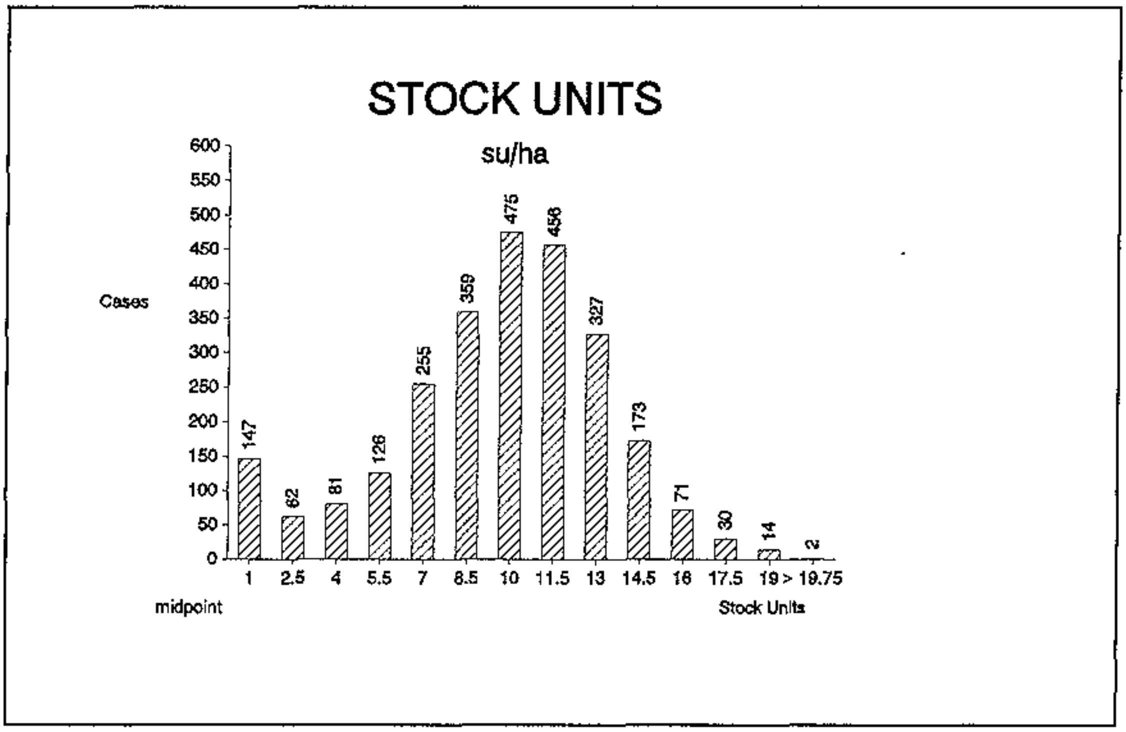

BEEF PRICE
Cases
400
350
300
250
200
150
100
50
0
4
15
52
157
215
262
286
290
184
123
81
36
13
6
7
122 182 242 302 362 422 482 542 602 662 722 782 842 902 > 930
midpoint
Dollars

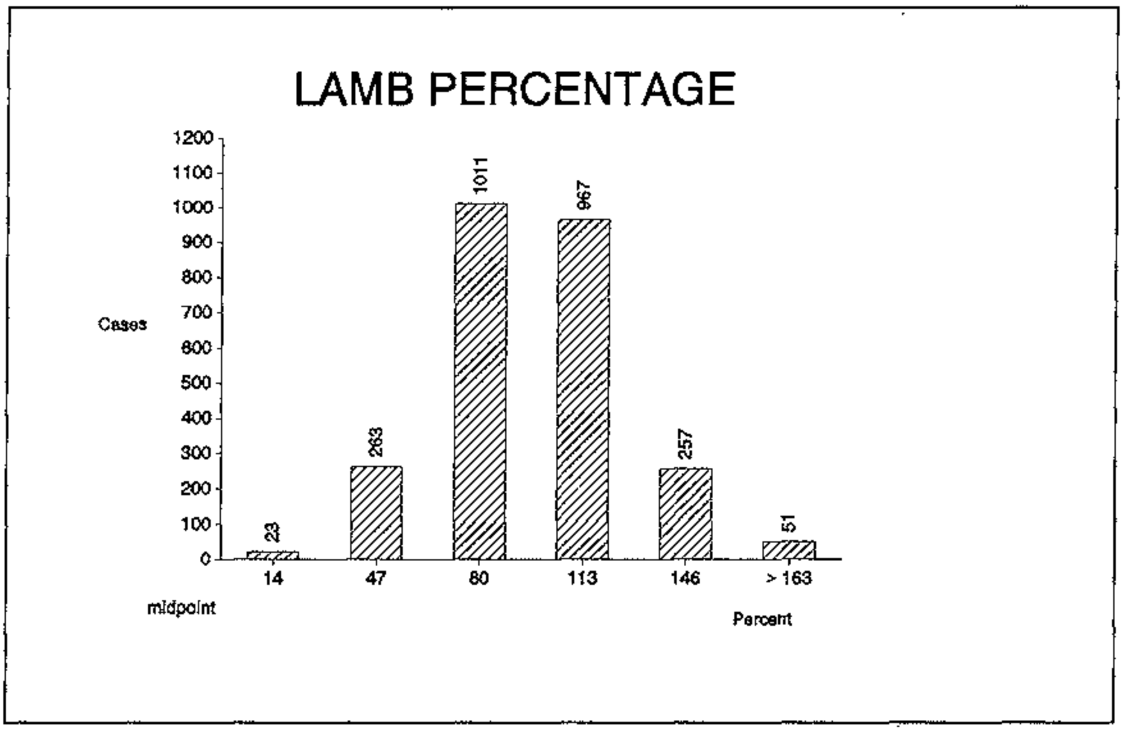

HEIFER PRICE
Cases
700
650
600
550
500
450
400
350
300
250
200
150
100
50
0
9
93
445
581
387
136
30
14
6
116 196 276 356 436 516 596 676 > 716
midpoint
Dollars

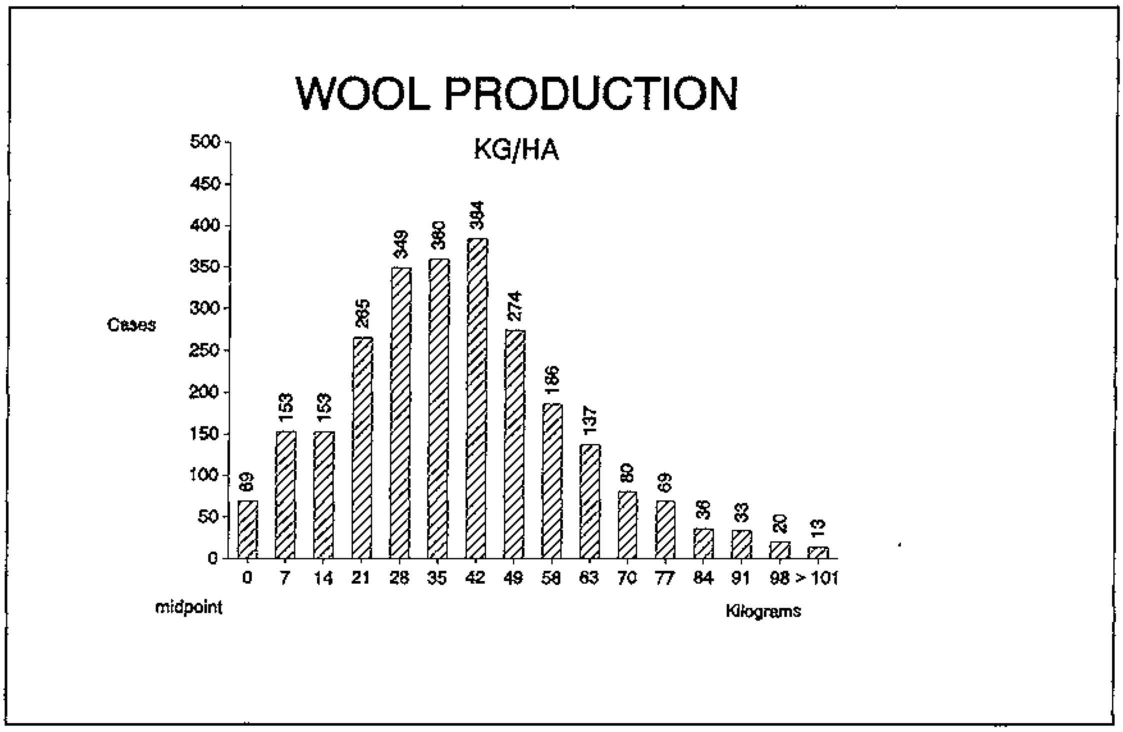

TOTAL EXPENSES
Cases
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
393
911
587
150
53
39
15
19
40
95
151
206
262
317
372
428
483
539
594
> 622
midpoint (thousands)
Dollars

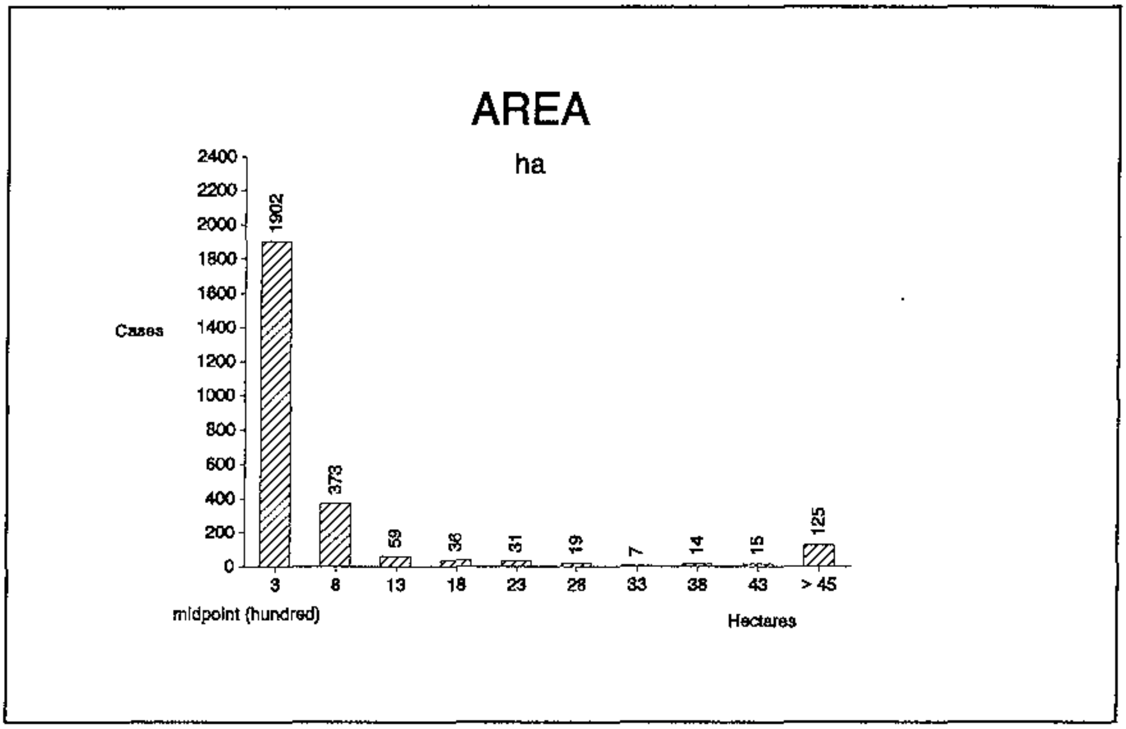

EQUITY %
Cases
900
800
700
600
500
400
300
200
100
0
35
23
31
79
710
< 0
0
12
24
36
48
60
72
84
96
midpoint
Percent


FERTILISER
kg/ha
Cases
550
500
450
400
350
300
250
200
150
100
50
0
454
376
291
288
212
167
118
50
33
19
16
0
75
150
225
300
375
450
525
600
675
> 725
midpoint
Kilograms


STOCK UNITS
su/ha
Cases
600
550
500
450
400
350
300
250
200
150
100
50
0
147
62
81
126
255
359
475
456
327
173
71
30
14
2
1
2.5
4
5.5
7
8.5
10
11.5
13
14.5
16
17.5
19
> 19.75
midpoint
Stock Units


LAMB PERCENTAGE
Cases
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
23
263
1011
967
257
51
14
47
80
113
146
> 163
midpoint
Percent


WOOL PRODUCTION
KG/HA
Cases
500
450
400
350
300
250
200
150
100
50
0
69
153
153
265
349
360
384
274
186
137
80
69
36
33
20
13
0
7
14
21
28
35
42
49
56
63
70
77
84
91
98
> 101
midpoint
Kilograms



AREA
ha
Cases
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
1902
373
59
36
31
19
7
14
15
125
3
8
13
18
23
28
33
38
43
> 45
midpoint (hundred)
Hectares

APPENDIX IV

COVARIATE MODEL RESULTS

MODEL 1

VAR.	CODE	COEFF	S.E	Z	L 95%	H 95%
CONST		2.43577	0.05452	44.677	2.3289	2.5426
CLASS	1	1.82789	0.06513	28.066	1.7002	1.9556
	2	0.23525	0.05992	3.926	0.1178	0.3527
	3	-2.06314	.	.	.	.
LOC.	1	-2.11256	0.06226	-33.929	-2.2346	-1.9905
	2	-2.23914	0.05886	-38.039	-2.3545	-2.1238
	3	-2.06829	0.05108	-40.492	-2.1684	-1.9682
	4	6.41999	.	.	.	.
C*L	1,1	-1.83017	0.09164	-19.971	-2.0098	-1.6506
	1,2	-1.91793	0.08522	-22.507	-2.0850	-1.7510
	1,3	-1.95120	.	.	.	.
	1,4	5.69930	.	.	.	.
	2,1	-0.22775	0.07798	- 2.921	-0.3806	-0.0749
	2,2	-0.17111	0.07836	- 2.184	-0.3247	-0.0175
	2,3	-0.27838	.	.	.	.
	2,4	0.67724	.	.	.	.
	3,1	2.05792	.	.	.	.
	3,2	2.08904	.	.	.	.
	3,3	2.22958	.	.	.	.
	3,4	-6.37654	.	.	.	.
WOOL		0.00000	.	.	.	.
LAMB		0.00000	.	.	.	.
EWE		0.00000	.	.	.	.
BEEF		0.00000	.	.	.	.
HEIF		0.00000	.	.	.	.
TTEXP		0.00000	.	.	.	.

FT/HA		0.00000	.	.	.	.
SU/HA		0.00000	.	.	.	.
LB %		0.00000	.	.	.	.
WL/HA		0.00000	.	.	.	.
AREA		0.00000	.	.	.	.

MODEL 2

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST		0.00000				
CLASS	1	-1.84822	0.10073	-18.348	-2.0457	-1.6508
	2	1.35359	0.07706	17.565	1.2026	1.5046
	3	0.49463	.	.	.	.
LOC.	1	0.23797	0.05529	4.304	0.0130	0.3464
	2	0.14461	0.05911	2.447	0.0288	0.2605
	3	0.59726	0.05715	10.451	0.4853	0.7093
	4	-0.97984	.	.	.	.
C*L	1,1	1.93597	0.12446	15.555	1.6920	2.1799
	1,2	1.82725	0.11318	16.145	1.6054	2.0491
	1,3	2.13512	0.09546	22.366	1.9480	2.3222
	1,4	-5.89834	.	.	.	.
	2,1	-1.31078	0.10220	-12.825	-1.5111	-1.1105
	2,2	-1.40223	0.10678	-13.132	-1.6115	-1.1929
	2,3	-1.66981	.	.	.	.
	2,4	4.38282	.	.	.	.
	3,1	-0.62519	.	.	.	.
	3,2	-0.42502	.	.	.	.
	3,3	-0.46531	.	.	.	.
	3,4	1.51552	.	.	.	.
LWOOL		18.24312	.	.	.	.
LLAMB		0.00000	.	.	.	.
LEWE		0.00000	.	.	.	.
LBEEF		0.78932	.	.	.	.

LHEIF		0.41361	.	.	.	.
LTTEXP		0.00000	.	.	.	.
LFT/HA		0.00000	.	.	.	.
LSU/HA		0.00000	.	.	.	.
LLB %		0.00000	.	.	.	.
LWL/HA		0.00000	.	.	.	.
LAREA		0.00000	.	.	.	.

LWOOL indicates base10 log of WOOL etc.

APPENDIX V

QUALITATIVE PRODUCTION VARIABLES

SIGNIFICANT INTERACTIONS

INTERACTIONS	D.F	PARTIAL X_2	P	ITER.
NTINC*CLASS*FT/HA*SU/HA	4	9.448	.0508	9
LOC. * CLASS * FT/HA	12	64.151	.0000	13
LOC. * CLASS * SU/HA	6	29.032	.0001	13
LOC. * FT/HA * SU/HA	6	17.859	.0066	14
CLASS * FT/HA * SU/HA	4	7.821	.0983	13
NTINC * LOC. * LB %	3	6.784	.0791	13
NTINC * CLASS * LB %	2	12.085	.0024	12
NTINC * FT/HA * LB %	2	17.058	.0002	13
CLASS* FT/HA * LB %	4	10.886	.0279	12
LOC. * CLASS * WL/HA	6	79.469	.0000	11
FT/HA * LB % * WL/HA	2	6.825	.0330	13
NTINC * LOC.	3	9.870	.0197	17
NTINC * CLASS	2	13.062	.0015	17
LOC. * CLASS	6	631.322	.0000	9
NTINC * FT/HA	2	7.207	.0272	17
LOC. * FT/HA	6	273.989	.0000	17
CLASS * FT/HA	4	81.324	.0000	16
NTINC * SU/HA	1	7.386	.0066	17
LOC. * SU/HA	3	527.354	.0000	10
CLASS * SU/HA	2	196.467	.0000	10
CLASS * LB %	2	24.174	.0000	17
LOC. * WL/HA	3	55.607	.0000	17
CLASS * WL/HA	2	103.690	.0000	18
FT/HA * WL/HA	2	9.840	.0073	17
SU/HA * WL/HA	1	374.549	.0000	19
LOC. * LB %	3	8.453	.0375	17

NTINC	1	265.283	.0000	2
LOC.	3	538.430	.0000	2
CLASS	2	239.848	.0000	2
FT/HA	2	990.569	.0000	2
SU/HA	1	29.023	.0000	2
LB %	1	3258.087	.0000	2

NONSIGNIFICANT INTERACTIONS

INTERACTIONS	D.F	PARTIAL X_2	P	ITER.
NTINC * LOC. * CLASS	6	7.267	.2968	14
NTINC * LOC. * FT/HA	6	5.665	.4617	13
NTINC * CLASS * FT/HA	4	2.104	.7166	14
NTINC * LOC. * SU/HA	3	3.932	.2689	12
NTINC * CLASS * SU/HA	2	.062	.9694	13
NTINC * FT/HA * SU/HA	2	.336	.8453	13
LOC. * CLASS * LB %	6	5.472	.4849	13
LOC. * FT/HA * LB %	6	4.410	.6214	12
NTINC * SU/HA * LB %	1	.020	.8875	14
LOC. * SU/HA * LB %	3	.461	.9274	13
CLASS * SU/HA * LB %	2	.179	.9144	14
FT/HA * SU/HA * LB %	2	.252	.8815	14
NTINC * LOC. * WL/HA	3	5.500	.1387	10
NTINC * CLASS * WL/HA	2	2.414	.2991	10
NTINC * FT/HA * WL/HA	2	1.203	.5480	13
LOC. * FT/HA * WL/HA	6	10.112	.1200	13
CLASS * FT/HA * WL/HA	4	1.293	.8626	13
NTINC * SU/HA * WL/HA	1	.054	.8165	13
LOC. * SU/HA * WL/HA	3	2.790	.4252	13
CLASS * SU/HA * WL/HA	2	3.005	.2226	13
FT/HA * SU/HA * WL/HA	2	.000	1.000	14
NTINC * LB % * WL/HA	1	2.143	.1432	13
LOC. * LB % * WL/HA	3	3.268	.3521	13

SU/HA * LB % * WL/HA	1	1.530	.2161	13
FT/HA * SU/HA	2	.533	.7662	17
NTINC * LB %	1	.273	.6014	17
FT/HA * LB %	2	.911	.6341	17
SU/HA * LB %	1	2.621	.1054	17
NTINC * WL/HA	1	.148	.7004	17
LB % * WL/HA	1	1.041	.3075	17
WL/HA	1	1.352	.2450	2

APPENDIX VI

QUALITATIVE PRODUCTION VARIABLES MODELS

MODEL 3

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST		1.63810	426.381	.004	-834.07	837.34
LOC.	1	.76292	330.465	.002	-646.94	648.47
	2	-4.08252	455.856	-0.009	-897.56	889.39
	3	-0.23281	0.109	-2.134	-0.45	-0.02
	4	3.55241	.	.	.	.
CLASS	1	3.52373	269.434	0.013	-524.56	531.62
	2	-4.50829	330.465	-0.014	-652.22	643.20
	3	0.98456	.	.	.	.
SU/HA	1	2.54520	0.105	24.237	2.34	2.75
	2	-2.54520	.	.	.	.
LB %	1	-1.56688	269.432	-0.006	-529.65	526.52
	2	1.56688	.	.	.	.
L*C	1,1	-0.90088	0.920	-0.098	-1.89	1.71
	1,2	1.12157	330.465	0.003	-646.59	648.83
	1,3	-0.22069	.	.	.	.
	2,1	-0.13752	0.879	-0.157	-1.86	1.59
	2,2	1.06773	330.465	0.003	-646.64	648.78
	2,3	-0.93021	.	.	.	.
	3,1	-0.49546	.	.	.	.
	3,2	-0.17391	.	.	.	.
	3,3	0.66937	.	.	.	.
	4,1	1.53386	.	.	.	.
	4,2	-2.01539	.	.	.	.
	4,3	0.48153	.	.	.	.
FT/HA	1	0.98701	0.117	8.441	0.76	1.22
	2	-0.64346	0.070	-9.192	-0.78	-0.51
	3	-0.34355	.	.	.	.
L*F	1,1	-0.89579	0.141	-6.361	-1.17	-0.62

	1,2	0.57927	0.082	7.071	0.42	0.74
	1,3	0.31652	.	.	.	.
	2,1	-0.85272	0.130	-6.565	-1.11	-0.60
	2,2	0.51425	0.078	6.569	0.36	0.67
	2,3	0.33847	.	.	.	.
	3,1	0.71227	0.152	4.693	0.42	1.01
	3,2	-0.38149	0.110	-3.461	-0.60	-0.17
	3,3	-0.33078	.	.	.	.
	4,1	1.03624	.	.	.	.
	4,2	-0.71203	.	.	.	.
	4,3	-0.32421	.	.	.	.
C*F	1,1	-9.6145	261.641	-0.037	-522.43	503.20
	1,2	2.05540	0.087	23.777	1.89	2.23
	1,3	7.5591	.	.	.	.
	2,1	8.33497	0.067	123.923	8.20	8.47
	2,2	-0.21827	0.063	-3.440	-0.34	-0.09
	2,3	-8.1167	.	.	.	.
	3,1	1.27953	.	.	.	.
	3,2	-1.83713	.	.	.	.
	3,3	0.5576	.	.	.	.
L*S	1,1	-2.4392	0.113	-21.595	-2.66	-2.22
	1,2	2.4392	.	.	.	.
	2,1	-2.49904	0.116	-21.533	-2.73	-2.27
	2,2	2.49904	.	.	.	.
	3,1	-1.00346	330.464	-0.003	-648.71	646.71
	3,2	1.00346	.	.	.	.
	4,1	5.9417	.	.	.	.
	4,2	-5.9417	.	.	.	.
C*S	1,1	2.20422	0.096	22.869	2.02	2.39
	1,2	-2.20422	.	.	.	.
	2,1	0.23333	0.084	2.788	0.07	0.40
	2,2	-0.23333	.	.	.	.
	3,1	-2.43755	.	.	.	.

	3,2	2.43755	.	.	.	.
L*L%	1,1	-0.34963	0.408	-0.856	-1.15	0.45
	1,2	0.34963	.	.	.	.
	2,1	4.47387	314.002	0.014	-610.97	619.92
	2,2	-4.47387	.	.	.	.
	3,1	0.06424	.	.	.	.
	3,2	-0.06424	.	.	.	.
	4,1	-4.18848	.	.	.	.
	4,2	4.18848	.	.	.	.
C*L%	1,1	-3.41436	269.432	-0.013	-531.50	524.67
	1,2	3.41436	.	.	.	.
	2,1	3.42192	.	.	.	.
	2,2	-3.42192	.	.	.	.
	3,1	-0.00756	.	.	.	.
	3,2	0.00756	.	.	.	.
L*C*S	1,1,1	-2.17592	0.107	-20.321	-2.39	-1.97
	1,1,2	2.17592	.	.	.	.
	1,2,1	-0.21776	0.101	-2.152	-0.42	-0.02
	1,2,2	0.21776	.	.	.	.
	1,3,1	2.39368	.	.	.	.
	1,3,2	-2.39368	.	.	.	.
	2,1,1	-2.41481	.	.	.	.
	2,1,2	2.41481	.	.	.	.
	2,2,1	-0.10868	.	.	.	.
	2,2,2	0.10868	.	.	.	.
	2,3,1	2.52349	.	.	.	.
	2,3,2	-2.52349	.	.	.	.
	3,1,1	0.00923	.	.	.	.
	3,1,2	-0.00923	.	.	.	.
	3,2,1	0.06261	330.460	0.0002	-647.65	647.77
	3,2,2	-0.06261	.	.	.	.
	3,3,1	-0.07184	.	.	.	.
	3,3,2	0.07184	.	.	.	.

	4,1,1	4.5815	.	.	.	.
	4,1,2	-4.5815	.	.	.	.
	4,2,1	0.26383	.	.	.	.
	4,2,2	-0.26383	.	.	.	.
	4,3,1	-4.84533	.	.	.	.
	4,3,2	4.84533	.	.	.	.
L*F*S	1,1,1	0.01893	0.089	0.212	-0.16	0.19
	1,1,2	-0.01893	.	.	.	.
	1,2,1	-0.00731	0.065	-0.113	-0.13	0.12
	1,2,2	0.00731	.	.	.	.
	1,3,1	-0.01162	.	.	.	.
	1,3,2	0.01162	.	.	.	.
	2,1,1	0.10366	0.079	1.310	-0.05	0.26
	2,1,2	-0.10366	.	.	.	.
	2,2,1	0.06796	0.075	0.906	-0.08	0.22
	2,2,2	-0.06796	.	.	.	.
	2,3,1	-0.17162	.	.	.	.
	2,3,2	0.17162	.	.	.	.
	3,1,1	-2.1161	0.187	-11.313	-2.48	-1.75
	3,1,2	2.1161	.	.	.	.
	3,2,1	1.01879	.	.	.	.
	3,2,2	-1.01879	.	.	.	.
	3,3,1	1.09731	.	.	.	.
	3,3,2	-1.09731	.	.	.	.
	4,1,1	1.99351	.	.	.	.
	4,1,2	-1.99351	.	.	.	.
	4,2,1	-1.07944	.	.	.	.
	4,2,2	1.07944	.	.	.	.
	4,3,1	-0.91407	.	.	.	.
	4,3,2	0.91407	.	.	.	.
C*F*L	1,1,1	9.65299	261.641	0.037	-503.16	522.47
	1,1,2	-9.65299	.	.	.	.
	1,2,1	-2.23884	.	.	.	.

	1,2,2	2.23884	.	.	.	.
	1,3,1	-7.41415	.	.	.	.
	1,3,2	7.41415	.	.	.	.
	2,1,1	-8.31619	.	.	.	.
	2,1,2	8.31619	.	.	.	.
	2,2,1	0.27737	.	.	.	.
	2,2,2	-0.27737	.	.	.	.
	2,3,1	8.03882	.	.	.	.
	2,3,2	-8.03882	.	.	.	.
	3,1,1	-1.3368	.	.	.	.
	3,1,2	1.3368	.	.	.	.
	3,2,1	1.96147	.	.	.	.
	3,2,2	-1.96147	.	.	.	.
	3,3,1	-0.62467	.	.	.	.
	3,3,2	0.62467	.	.	.	.
WOOL		1.49774	0.703	2.132	0.12	2.88
TTEXP		-.000002	0.000	-0.538	-.00001	.00001
AREA		0.00281	0.001	2.266	0.0004	0.005

MODEL 4

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST		0.00000				
LOC.	1	0.24755	0.05159	4.7989	0.1465	0.3487
	2	-0.01694	0.05836	-0.2903	-0.1313	0.0974
	3	-0.62086	0.19323	-3.2130	-0.9996	-0.2421
	4	0.39025	.	.	.	.
FT/HA	1	-0.65636	0.14957	-4.3884	-0.9495	-0.3632
	2	0.20582	0.06593	3.1220	0.0766	0.3350
	3	0.45054	.	.	.	.
SU/HA	1	0.27416	0.09291	2.9509	0.0921	0.4563
	2	-0.27416	.	.	.	.
L*F	1,1	0.70885	0.15505	4.5718	0.4050	1.0127

	1,2	-0.34634	0.10383	-3.3357	-0.5498	-0.1428
	1,3	-0.36251	.	.	.	.
	2,1	0.78800	0.14010	5.6245	0.5134	1.0626
	2,2	-0.34582	0.10551	-3.2778	-0.5526	-0.1390
	2,3	-0.44218	.	.	.	.
	3,1	-2.13280	0.46203	-4.6161	-3.0384	-1.2272
	3,2	1.45998	0.37734	3.8692	0.7204	2.1996
	3,3	0.67282	.	.	.	.
	4,1	0.63595	.	.	.	.
	4,2	-0.76780	.	.	.	.
	4,3	0.13185	.	.	.	.
L*S	1,1	0.03182	0.06296	0.5054	-0.0916	0.1552
	1,2	-0.03182	.	.	.	.
	2,1	-0.02050	0.06099	-0.3361	-0.1400	0.0990
	2,2	0.02050	.	.	.	.
	3,1	0.22435	0.11689	1.9194	-0.0048	0.4534
	3,2	-0.22435	.	.	.	.
	4,1	-0.23567	.	.	.	.
	4,2	0.23567	.	.	.	.
WOOL		4.79321	1.28347	3.7346	2.2776	7.3088
TTEXP		0.00002	0.00000	6.3663	.00002	.00003
AREA		-0.00063	0.00056	-1.1190	-0.0017	0.0005

MODEL 5

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
LOC.	1	0.19131	0.04359	4.3891	0.1059	0.2767
	2	-0.06806	0.05194	-1.3104	-0.1699	0.0337
	3	-0.61172	0.16209	-3.7739	-0.9294	-0.2940
	4	0.48847	.	.	.	.
FT/HA	1	-0.58701	0.12265	-4.7858	-0.8274	-0.3466
	2	0.16094	0.04344	3.7052	0.0758	0.2461
	3	0.42607	.	.	.	.

SU/HA	1	0.25264	0.05452	4.6339	0.1458	0.3595
	2	-0.25264	.	.	.	.
L*F	1,1	0.64828	0.13611	4.7631	0.3815	0.9151
	1,2	-0.28674	0.08363	-3.4285	-0.4507	-0.1228
	1,3	-0.36154	.	.	.	.
	2,1	0.71656	0.11059	6.4796	0.4998	0.9333
	2,2	-0.26559	0.08043	-3.3022	-0.4232	-0.1079
	2,3	-0.45097	.	.	.	.
	3,1	-1.90245	0.37706	-5.0454	-2.6415	-1.1634
	3,2	1.15986	0.28901	4.0132	0.5934	1.7263
	3,3	0.74259	.	.	.	.
	4,1	0.53761	.	.	.	.
	4,2	-0.60753	.	.	.	.
	4,3	0.06992	.	.	.	.
WOOL		4.18089	1.09462	3.8195	2.0354	6.3263
TTEXP		0.00002	0.00000	7.7679	0.00002	0.00003
AREA		-0.00120	0.00042	-2.9096	-0.0020	-0.0004

APPENDIX VII

ALL VARIABLES QUALITATIVE INTERACTIONS

SIGNIFICANT INTERACTIONS

INTERACTIONS	DF	X_2	P	ITER.
LOC. * SU/HA * FT/HA	2	7.570	0.0227	12
LOC. * SU/HA * WOOL	2	11.548	0.0031	12
LOC. * FT/HA * TTEXP	4	26.165	0.0000	12
FT/HA * SU/HA * TTEXP	4	9.317	0.0537	11
NTINC * FT/HA * AREA	4	16.423	0.0025	9
LOC. * FT/HA * AREA	4	13.211	0.0103	12
LOC. * SU/HA * AREA	2	5.638	0.0597	12
FT/HA * SU/HA * AREA	4	18.641	0.0009	12
NTINC * TTEXP * AREA	4	12.927	0.0116	13
LOC. * TTEXP * AREA	4	21.760	0.0002	13
SU/HA * TTEXP * AREA	4	12.093	0.0167	12
NTINC * LOC.	1	3.854	0.0496	15
NTINC * FT/HA	2	20.397	0.0000	15
LOC. * FT/HA	2	163.273	0.0000	15
NTINC * SU/HA	1	15.631	0.0001	15
LOC. * SU/HA	1	181.573	0.0000	14
FT/HA * SU/HA	2	9.710	0.0078	15
NTINC * WOOL	2	4.887	0.0869	15
LOC. * WOOL	2	129.013	0.0000	14
FT/HA * WOOL	4	13.405	0.0095	16
SU/HA * WOOL	2	12.559	0.0019	16
NTINC * TTEXP	2	202.842	0.0000	13
LOC. * TTEXP	2	27.708	0.0000	14
FT/HA * TTEXP	4	130.454	0.0000	15
SU/HA * TTEXP	2	17.846	0.0001	13
WOOL * TTEXP	4	7.817	0.0985	15

NTINC * AREA	2	140.229	0.0000	13
LOC. * AREA	2	93.473	0.0000	14
FT/HA * AREA	4	136.757	0.0000	14
SU/HA * AREA	2	328.801	0.0000	12
WOOL * AREA	4	114.015	0.0000	15
TTEXP * AREA	4	839.848	0.0000	9
NTINC	1	262.567	0.0000	2
LOC.	1	123.405	0.0000	2
FT/HA	2	999.767	0.0000	2
SU/HA	1	27.060	0.0000	2
WOOL	2	4248.446	0.0000	2
TTEXP	2	366.247	0.0000	2
AREA	2	534.998	0.0000	2

NONSIGNIFICANT INTERACTIONS

INTERACTIONS	DF	X_2	P	ITER.
NTINC * LOC. * FT/HA	2	1.170	0.5571	12
NTINC * LOC. * SU/HA	1	0.053	0.8177	12
NTINC * FT/HA * SU/HA	2	0.817	0.6648	12
NTINC * LOC. * WOOL	2	0.002	0.9991	12
NTINC * FT/HA * WOOL	4	1.031	0.9050	12
LOC. * FT/HA * WOOL	4	0.783	0.9407	12
NTINC * SU/HA * WOOL	2	1.888	0.3891	12
FT/HA * SU/HA * WOOL	4	2.477	0.6487	12
NTINC * LOC. * TTEXP	2	0.692	0.7075	12
NTINC * FT/HA * TTEXP	4	6.398	0.1713	9
NTINC * SU/HA * TTEXP	2	2.856	0.2397	12
LOC. * SU/HA * TTEXP	2	1.186	0.5526	12
NTINC * WOOL * TTEXP	4	1.411	0.8423	12

LOC. * WOOL * TTEXP	4	2.971	0.5627	12
FT/HA * WOOL * TTEXP	8	3.399	0.9069	12
SU/HA * WOOL * TTEXP	4	3.150	0.5331	12
NTINC * LOC. * AREA	2	1.968	0.3738	12
NTINC * SU/HA * AREA	2	0.340	0.8437	11
NTINC * WOOL * AREA	4	0.261	0.9922	12
LOC. * WOOL * AREA	4	0.000	1.0000	12
FT/HA * WOOL * AREA	8	2.434	0.9647	12
SU/HA * WOOL * AREA	4	0.037	0.9998	12
FT/HA * TTEXP * AREA	8	4.016	0.8557	13
WOOL * TTEXP * AREA	8	4.619	0.7974	12

APPENDIX VIII

MODEL 6

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST.		4.23983	0.131	32.327	3.98	4.50
AREA	1	2.02887	266.647	0.008	-520.60	524.66
	2	-1.40424	0.130	-10.786	-1.66	-1.15
	3	-0.62463	.	.	.	.
TTEXP	1	-1.89607	0.122	-15.587	-2.14	-1.66
	2	0.04207	0.095	0.442	-0.15	0.23
	3	1.85400	.	.	.	.
WOOL	1	5.27831	0.253	20.870	4.78	5.77
	2	-2.51405	.	.	.	.
	3	-2.76426	.	.	.	.
SU/HA	1	-1.32939	0.123	-10.854	-1.57	-1.09
	2	1.32939	.	.	.	.
LOC.	1	-0.12735	0.097	-1.320	-0.32	0.06
	2	0.12735	.	.	.	.
E*A	1,1	-1.40036	0.189	-7.421	-1.77	-1.03
	1,2	1.53460	0.121	12.686	1.30	1.77
	1,3	-0.13424	.	.	.	.
	2,1	-2.82256	0.160	-17.628	-3.14	-2.51
	2,2	-0.14805	0.111	-1.338	-0.37	0.07
	2,3	2.97061	.	.	.	.
	3,1	-1.57025	.	.	.	.
	3,2	-1.38655	.	.	.	.
	3,3	-2.83637	.	.	.	.
W*A	1,1	-0.28446	533.295	-0.001	-1045.5	1044.9
	1,2	0.10433	0.216	0.482	-0.32	0.53
	1,3	0.18013	.	.	.	.

	2,1	0.01805	266.647	0.00007	-522.61	522.65
	2,2	-0.00041	.	.	.	.
	2,3	-0.01764	.	.	.	.
	3,1	0.26641	.	.	.	.
	3,2	-0.10392	.	.	.	.
	3,3	-0.16249	.	.	.	.
S*A	1,1	0.02190	0.089	0.247	-0.15	0.19
	1,2	0.07328	0.086	0.856	-0.09	0.24
	1,3	-0.09518	.	.	.	.
	2,1	0.16249	.	.	.	.
	2,2	-0.02190	.	.	.	.
	2,3	0.09518	.	.	.	.
W*E	1,1	0.27060	0.208	1.301	-0.14	0.68
	1,2	-0.06041	0.157	-0.384	-0.37	0.25
	1,3	-0.21019	.	.	.	.
	2,1	-0.16543	.	.	.	.
	2,2	0.16341	.	.	.	.
	2,3	0.00202	.	.	.	.
	3,1	-0.10517	.	.	.	.
	3,2	-0.10300	.	.	.	.
	3,3	0.20817	.	.	.	.
S*E	1,1	0.04114	0.057	0.722	-0.07	0.15
	1,2	-0.00416	0.052	-0.080	-0.11	0.10
	1,3	-0.03698	.	.	.	.
	2,1	-0.04114	.	.	.	.
	2,2	0.00416	.	.	.	.
	2,3	0.03698	.	.	.	.
L*E	1,1	-1.43948	0.080	-18.015	-1.60	-1.28
	1,2	0.02401	0.062	0.389	-0.10	0.14
	1,3	1.41547	.	.	.	.
	2,1	1.43948	.	.	.	.
	2,2	-0.02401	.	.	.	.

	2,3	-1.41547	.	.	.	.
S*W	1,1	-0.43709	0.382	-1.145	-1.19	0.31
	1,2	-0.00375	.	.	.	.
	1,3	0.44084	.	.	.	.
	2,1	0.43709	.	.	.	.
	2,2	0.00375	.	.	.	.
	2,3	-0.44084	.	.	.	.
L*W	1,1	-0.05773	0.409	-0.141	-0.86	0.74
	1,2	-0.02032	.	.	.	.
	1,3	0.07805	.	.	.	.
	2,1	0.05773	.	.	.	.
	2,2	0.02032	.	.	.	.
	2,3	-0.07805	.	.	.	.
L*S	1,1	-0.08559	0.049	-1.748	-0.18	0.01
	1,2	0.08559	.	.	.	.
	2,1	0.08559	.	.	.	.
	2,2	-0.08559	.	.	.	.
S*E*A	1,1,1	1.41491	0.160	8.834	1.10	1.73
	1,1,2	0.08592	.	.	.	.
	1,1,3	-1.50083	.	.	.	.
	1,2,1	1.47018	.	.	.	.
	1,2,2	0.00592	.	.	.	.
	1,2,3	-1.47610	.	.	.	.
	1,3,1	-2.88509	.	.	.	.
	1,3,2	-0.09184	.	.	.	.
	1,3,3	2.97693	.	.	.	.
	2,1,1	-1.41491	.	.	.	.
	2,1,2	-0.08592	.	.	.	.
	2,1,3	1.50083	.	.	.	.
	2,2,1	-1.47018	.	.	.	.
	2,2,2	-0.00592	.	.	.	.
	2,2,3	1.47610	.	.	.	.
	2,3,1	2.88509	.	.	.	.

	2,3,2	0.09184	.	.	.	.
	2,3,3	-2.97693	.	.	.	.
L*E*A	1,1,1	0.10313	0.130	0.792	-0.15	0.36
	1,1,2	1.62237	0.086	18.930	1.45	1.79
	1,1,3	-1.72550	.	.	.	.
	1,2,1	-1.29959	.	.	.	.
	1,2,2	-0.05599	.	.	.	.
	1,2,3	1.35558	.	.	.	.
	1,3,1	1.19646	.	.	.	.
	1,3,2	-1.56638	.	.	.	.
	1,3,3	0.36992	.	.	.	.
	2,1,1	-0.10313	.	.	.	.
	2,1,2	-1.62237	.	.	.	.
	2,1,3	1.72550	.	.	.	.
	2,2,1	1.29959	.	.	.	.
	2,2,2	0.05599	.	.	.	.
	2,2,3	-1.35558	.	.	.	.
	2,3,1	-1.19646	.	.	.	.
	2,3,2	1.56638	.	.	.	.
	2,3,3	-0.36992	.	.	.	.
L*S*A	1,1,1	0.08928	0.065	1.377	-0.04	0.22
	1,1,2	-0.04514	.	.	.	.
	1,1,3	-0.04414	.	.	.	.
	1,2,1	-0.08928	.	.	.	.
	1,2,2	0.04514	.	.	.	.
	1,2,3	0.04414	.	.	.	.
	2,1,1	-0.08928	.	.	.	.
	2,1,2	0.04514	.	.	.	.
	2,1,3	0.04414	.	.	.	.
	2,2,1	0.08928	.	.	.	.
	2,2,2	-0.04514	.	.	.	.
	2,2,3	-0.04414	.	.	.	.
L*S*W	1,1,1	-0.01269	.	.	.	.

	1,1,2	0.04502	.	.	.	.
	1,1,3	-0.03233	.	.	.	.
	1,2,1	0.01269	.	.	.	.
	1,2,2	-0.04502	.	.	.	.
	1,2,3	0.03233	.	.	.	.
	2,1,1	0.01269	.	.	.	.
	2,1,2	-0.04502	.	.	.	.
	2,1,3	0.03233	.	.	.	.
	2,2,1	-0.01269	.	.	.	.
	2,2,2	0.04502	.	.	.	.
	2,2,3	-0.03233	.	.	.	.

MODEL 7

VAR.	CODE	COEFF.	S.E	Z	L 95%	H 95%
CONST		3.08473	0.06566	46.978	2.956	3.213
LOC.	1	-0.03782	0.02575	-1.469	-0.088	0.013
	2	0.03782	.	.	.	.
SU/HA	1	0.11128	0.02652	4.196	0.059	0.163
	2	-0.11128	.	.	.	.
WOOL	1	5.19743	0.09848	52.774	5.004	5.390
	2	-2.63156	.	.	.	.
	3	-2.56587	.	.	.	.
TTEXP	1	-0.57121	0.07610	-7.506	-0.720	-0.422
	2	0.06073	0.06999	0.868	-0.076	0.198
	3	0.51048	.	.	.	.
AREA	1	0.66329	0.11962	5.545	0.429	0.898
	2	-0.11171	0.06959	-1.605	-0.248	0.025
	3	-0.55158	.	.	.	.
T*A	1,1	-0.24432	0.12571	-1.943	-0.491	0.002
	1,2	0.14109	0.08184	1.724	-0.019	0.302
	1,3	0.10323	.	.	.	.

	2,1	-0.05055	0.12520	-0.404	-0.296	0.195
	2,2	-0.04596	0.07550	-0.609	-0.194	0.102
	2,3	0.09651	.	.	.	.
	3,1	0.29487	.	.	.	.
	3,2	-0.09513	.	.	.	.
	3,3	-0.19974	.	.	.	.

APPENDIX IX

MICRO MODEL CROSSTABULATION

WOOL MODEL CROSSTABULATION

NTINC				1					
	AREA		1	LOC		2			SAMP
WOOL		1	2	3	1	2	3	TOT	TOT
TTEXP	1	120	209	12	49	71	14	475	848
1	2	4	3	1	2	3	0	13	45
	3	0	0	0	1	3	1	5	18
	TOT	124	212	13	52	77	15	493	911
	1	53	167	21	39	87	8	375	485
2	2	20	33	5	6	15	6	85	162
	3	3	5	0	6	17	15	46	86
	TOT	76	205	26	51	119	29	506	733
	1	19	102	18	44	80	12	275	306
3	2	23	70	15	13	39	9	169	238
	3	26	57	5	15	56	97	256	392
	TOT	68	229	38	72	175	118	700	936
TOTAL		268	646	77	175	371	162	1699	2580

NTINC				2					
	AREA		1	LOC		2			SAMP
WOOL		1	2	3	1	2	3	TOT	TOT
TTEXP	1	80	182	12	22	62	15	**373**	**848**
1	2	6	15	0	3	6	2	**32**	**45**
	3	0	1	0	1	6	5	**13**	**18**
	TOT	**86**	**198**	**12**	**26**	**74**	**22**	**418**	**911**
	1	17	55	9	7	21	1	**110**	**485**
2	2	13	47	3	1	9	4	**77**	**162**
	3	5	5	2	1	16	11	**40**	**86**
	TOT	**35**	**107**	**14**	**9**	**46**	**16**	**227**	**733**
	1	1	10	2	4	13	1	**31**	**306**
3	2	7	36	10	3	9	4	**69**	**238**
	3	11	46	5	4	27	43	**136**	**392**
	TOT	**19**	**91**	**17**	**11**	**49**	**48**	**236**	**936**
TOTAL		**140**	**397**	**43**	**46**	**169**	**86**	**881**	**2580**

LAMB MODEL CROSSTABULATION

NTINC				1					
	AREA		1	LOC.		2			SAMP
LAMB		1	2	3	1	2	3	TOT	TOT
TTEXP	1	173	116	41	53	59	22	**464**	**826**
1	2	5	3	0	4	1	0	**13**	**45**
	3	0	0	0	2	2	0	**4**	**14**
	TOT	**178**	**119**	**41**	**59**	**62**	**22**	**481**	**885**
	1	93	97	41	42	57	35	**365**	**473**
2	2	32	19	2	15	10	2	**80**	**157**
	3	6	2	0	24	13	0	**45**	**83**
	TOT	**131**	**118**	**43**	**81**	**80**	**37**	**490**	**713**
	1	41	72	25	32	68	36	**274**	**305**
3	2	49	40	19	14	29	18	**169**	**238**
	3	36	40	12	71	67	22	**248**	**379**
	TOT	**126**	**152**	**56**	**117**	**164**	**76**	**691**	**922**
TOTAL		**435**	**389**	**140**	**257**	**306**	**135**	**1662**	**2520**

NTINC				2					
	AREA		1	LOC.		2			SAMP
LAMB		1	2	3	1	2	3	TOT	TOT
TTEXP	1	91	136	41	24	49	21	362	826
1	2	10	8	3	5	6	0	32	45
	3	0	0	1	2	6	1	10	14
	TOT	101	144	45	31	61	22	404	885
	1	18	36	25	5	16	8	108	473
2	2	22	30	11	7	6	1	77	157
	3	7	4	1	15	10	1	38	83
	TOT	47	70	37	27	32	10	223	713
	1	1	7	5	1	10	7	31	305
3	2	18	22	13	4	7	5	69	238
	3	22	24	16	22	30	17	131	379
	TOT	41	53	34	27	47	29	231	922
TOTAL		189	267	116	85	140	61	858	2520

EWE MODEL CROSSTABULATION

NTINC				1			
	AREA		1	LOC.	2		SAMP
EWE		1	2	1	2	TOT	TOT
TTEXP	1	235	90	96	34	455	822
1	2	5	2	4	1	12	43
	3	0	0	4	1	5	18
	TOT	240	92	104	36	472	883
	1	146	91	82	49	368	478
2	2	43	15	20	6	84	161
	3	7	1	28	9	45	85
	TOT	196	107	130	64	497	724
	1	80	59	76	58	273	304
3	2	68	40	34	26	168	237
	3	59	29	72	89	249	383
	TOT	207	128	182	173	690	924
TOTAL		643	327	416	273	1659	2531

NTINC				2			
	AREA		1	LOC.	2		SAMP
EWE		1	2	1	2	TOT	TOT
TTEXP	1	190	81	61	35	367	822
1	2	13	7	4	7	31	43
	3	0	1	5	7	13	18
	TOT	203	89	70	49	411	883
	1	46	35	16	13	110	478
2	2	45	18	10	4	77	161
	3	7	5	14	14	40	85
	TOT	98	58	40	31	227	724
	1	5	8	10	8	31	304
3	2	31	22	7	9	69	237
	3	34	28	25	47	134	383
	TOT	70	58	42	64	234	924
TOTAL		371	205	152	144	872	2531

BEEF MODEL CROSSTABULATION

NTINC				1					
	AREA		1	LOC.		2			SAMP
BEEF		1	2	3	1	2	3	TOT	TOT
TTEXP	1	108	72	45	13	11	8	257	494
1	2	5	0	1	1	1	1	9	34
	3	0	0	0	1	1	0	2	10
	TOT	113	72	46	15	13	9	268	538
	1	83	59	47	15	23	4	231	313
2	2	25	13	6	4	10	2	60	132
	3	6	1	1	14	10	2	34	70
	TOT	114	73	54	33	43	8	325	515
	1	45	40	31	13	18	9	156	173
3	2	33	35	23	7	20	10	128	183
	3	39	35	11	57	50	20	212	322
	TOT	117	110	65	77	88	39	496	678
TOTAL		344	255	165	125	144	56	1089	1731

NTINC				2					
	AREA		1	LOC.		2			SAMP
BEEF		1	2	3	1	2	3	TOT	TOT
TTEXP	1	77	87	41	9	17	6	237	494
1	2	14	3	0	1	3	4	25	34
	3	0	0	1	3	2	2	8	10
	TOT	91	90	42	13	22	12	270	538
	1	19	30	17	4	8	4	82	313
2	2	24	29	9	3	6	1	72	132
	3	5	6	1	11	11	2	36	70
	TOT	48	65	27	18	25	7	190	515
	1	2	3	7	2	2	1	17	173
3	2	16	13	19	1	6	0	55	183
	3	23	25	12	18	21	11	110	322
	TOT	41	41	38	21	29	12	182	678
TOTAL		180	196	107	52	76	31	642	1731

HEIFER MODEL CROSSTABULATION

NTINC			1				
	AREA	1	LOC.	2			SAMP
HEIFER		1	2	1	2	TOT	TOT
TTEXP	1	74	151	7	35	267	501
1	2	1	3	1	1	6	29
	3	0	0	1	2	3	13
	TOT	75	154	9	38	276	543
	1	47	113	8	29	197	271
2	2	16	29	4	15	64	128
	3	6	2	6	22	36	75
	TOT	69	144	18	66	297	474
	1	22	85	5	28	140	158
3	2	17	76	2	33	128	180
	3	26	60	50	91	227	346
	TOT	65	221	57	152	495	684
TOTAL		209	519	84	256	1068	1701

NTINC			2				
	AREA	**1**	**LOC.**	**2**			**SAMP**
HEIFER		**1**	**2**	**1**	**2**	**TOT**	**TOT**
TTEXP	**1**	49	155	5	25	234	
1	**2**	7	11	2	3	23	29
	3	0	1	5	4	10	13
	TOT	**56**	**167**	**12**	**32**	**267**	**543**
	1	11	48	4	11	74	271
2	**2**	15	42	1	6	64	128
	3	5	7	4	23	39	75
	TOT	**31**	**97**	**9**	**40**	**177**	**474**
	1	1	10	1	6	18	158
3	**2**	8	35	0	9	52	180
	3	10	48	23	38	119	346
	TOT	**19**	**93**	**24**	**53**	**189**	**684**
TOTAL		**106**	**357**	**45**	**125**	**633**	**1701**

APPENDIX X

MODEL RESIDUAL PLOTS

WOOL MODEL

Observed counts VS Adjusted residuals

```
       --+----+----+----+----+----+----+----+----+--
     2 +                                  1        +
A      | 11       1       1                        |
D      | 1112    11                                |
J      | 1414              1                       |
       | 2824  1     1             1               |
R    0 + 554 41   11   1                           +
S      | 55211  2  1 1      1                      |
D      | 3421   1                                  |
       | 2 2      11  1                            |
       | 2 1   1     1                             |
    -2 +                                1          +
      ---+----+----+----+----+----+----+----+----+--
         0   30   60   90  120  150  180  210  240

                   OBSERVED COUNT
```

Normal Plot

```
            --+----+----+----+----+----+----+----+----+-
      2.5 +                                        1 +
E         |                                      2 1 |
X         |                                  1141    |
P         |                            3322111       |
          |                        28171             |
V       0 +                    33B33                 +
A         |                17182                     |
L         |          1112233                         |
U         |       1411                               |
E         |   1 2                                    |
     -2.5 +   1                                      +
          --+----+----+----+----+----+----+----+----+-
               -1.5      -.5       .5       1.5
          -2        -1        0         1         2

                   ADJUSTED RESIDUAL
```

LAMB MODEL

```
              Observed counts VS Adjusted residuals

         -+----+----+----+----+----+----+----+----+-
       2 +        1                                +
A        |  1 1                                    |
D        |  11111 2      1           1       1     |
J        |44111111  1   1                          |
         |32 2112 1                                |
R      0 +761251 11           2                    +
S        |33112      1   1                         |
D        |61211 11  111                            |
         |131  1  1      1  1     1                |
         |1      1                                 |
      -2 +        1                                +
         -+----+----+----+----+----+----+----+----+-
          0   25   50   75   100  125  150  175  200

                         OBSERVED COUNT
```

```
                         Normal Plot

         --+----+----+----+----+----+----+----+----+-
     2.5 +                                       1 +
E        |                                  1  2   |
X        |                                331      |
P        |                            25141        |
         |                       431443            |
V      0 +                   14D41                 +
A        |              344134                     |
L        |          14152                          |
U        |        133                              |
E        |     2  1                                |
    -2.5 +   1                                     +
         --+----+----+----+----+----+----+----+----+-
              -1.5       -.5        .5        1.5
          -2        -1         0          1         2

                        ADJUSTED RESIDUAL
```

EWE MODEL

```
              Observed counts VS Adjusted residuals

         -+----+----+----+----+----+----+----+----+-
       2 +      1         1                          +
A        |                1                          |
D        | 31    11                                  |
J        | 211   11 2 2                              |
         |13  1  1     1                             |
R      0 +32231 112       12             1       1   +
S        |121 1    2                                 |
D        |132    1  1     1      1                   |
         |21    11     1                             |
         | 1                                         |
      -2 +  1           1                            +
         -+----+----+----+----+----+----+----+----+-
          0   30   60   90  120  150  180  210  240

                      OBSERVED COUNT
```

```
                          Normal Plot

        --+----+----+----+----+----+----+----+----+-
   2.5 +                                          1 +
E      |                                        11  |
X      |                                  211       |
P      |                              2232          |
       |                        21133 3             |
V    0 +                      383                   +
A      |                3 33112                     |
L      |             2322                           |
U      |          112                               |
E      |     11                                     |
  -2.5 +   1                                        +
       --+----+----+----+----+----+----+----+----+-
             -1.5       -.5        .5       1.5
          -2        -1          0        1        2

                        ADJUSTED RESIDUAL
```

BEEF MODEL

Observed counts VS Adjusted residuals

```
         -+----+----+----+----+----+----+----+----+-
       2 +                                          +
A        | 11 1 1                                   |
D        |      1   1 11  1                         |
J        | 23211  1 1   1 1            1            |
         |451211 1     1                       1    |
R      0 +7533 12 3   1     1     1    1            +
S        |721 21111                 1               |
D        |212 3 11   11 1     1                     |
         | 1  2   1     1                           |
         |21  1                                     |
      -2 +                                          +
         -+----+----+----+----+----+----+----+----+-
          0   15   30   45   60   75   90  105   120

                   OBSERVED COUNT
```

Normal Plot

```
          --+----+----+----+----+----+----+----+----+-
      2.5 +                                     1    +
E         |                                    21    |
X         |                               2221       |
P         |                           25132          |
          |                       43732              |
V       0 +                    787                   +
A         |               23734                      |
L         |           23152                          |
U         |        1222                              |
E         |    12                                    |
      2.5 +    1                                     +
          --+----+----+----+----+----+----+----+----+-
               -1.8      -.6       .6        1.8
          -2.4      -1.2       0        1.2       2.4

                     ADJUSTED RESIDUAL
```

HEIFER MODEL

```
              Observed counts VS Adjusted residuals

            -+----+----+----+----+----+----+----+----+-
          2 +                                          +
    A       | 11  1                                    |
    D       |  1      1    1       1                   |
    J       |1331     211 11     1   1                 |
            |1311          1     1                     |
    R     0 +31  2 1  1                          11    +
    S       |231      1                  1             |
    D       |43121 12  1 1                             |
            |1  1         1 1                          |
            |  2    1                                  |
         -2 +                                          +
            -+----+----+----+----+----+----+----+----+-
             0   20   40   60   80   100  120  140  160

                          OBSERVED COUNT
```

```
                          Normal Plot

            --+----+----+----+----+----+----+----+----+-
        2.5 +                                       1  +
    E       |                                     2    |
    X       |                                1111      |
    P       |                              162         |
            |                         113224           |
    V     0 +                   13 6 31                +
    A       |              422311                      |
    L       |            261                           |
    U       |         1111                             |
    E       |     2                                    |
       -2.5 +   1                                      +
            --+----+----+----+----+----+----+----+----+-
                 -1.5       -.5        .5        1.5
             -2        -1         0          1         2

                           ADJUSTED RESIDUAL
```

APPENDIX XI

MICRO MODEL RISK COMBINATIONS

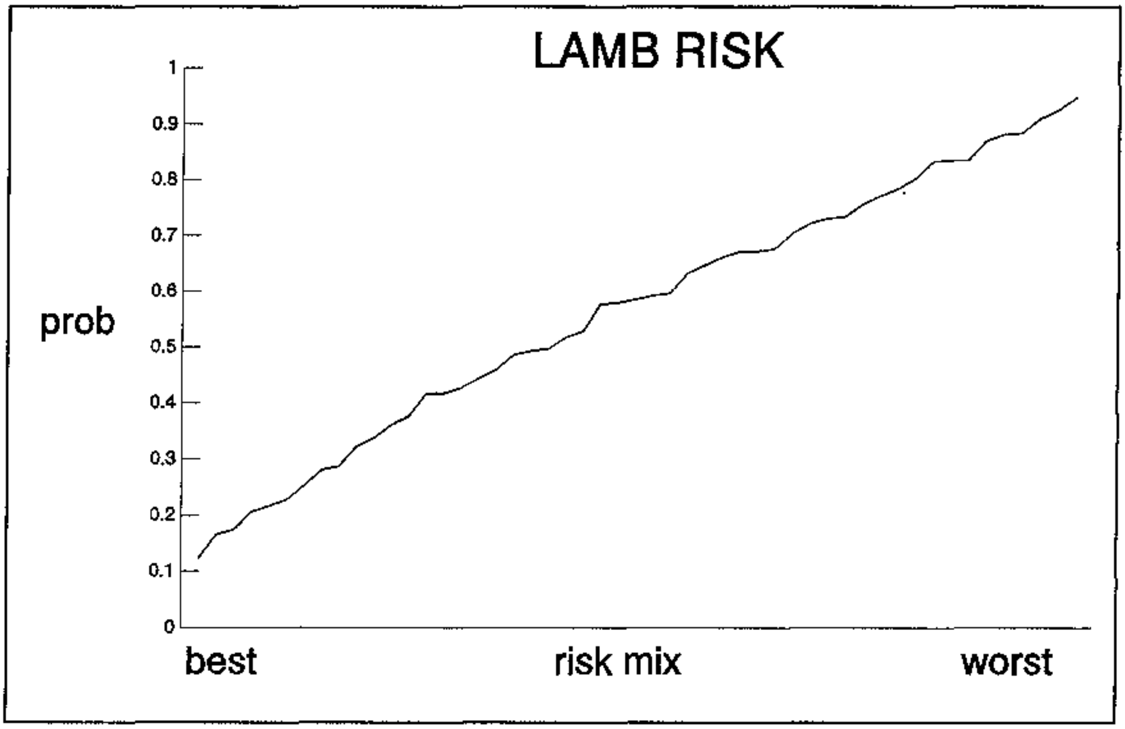
WOOL		LAMB		EWE		BEEF		HEIF	
PIAE	RISK	PIAE	RISK	PIAE	RISK	PIAE	RISK	PIAE	RISK
3131	.194	3131	.123	2121	.219	3131	.149	1131	.234
3121	.240	3121	.165	1131	.226	2121	.171	2231	.241
2121	.246	3231	.174	2231	.229	3121	.193	2121	.243
3231	.251	2121	.206	1121	.280	3231	.206	1231	.305
2231	.257	2231	.217	2221	.284	2231	.208	1121	.307
1131	.271	3221	.228	1231	.292	1131	.217	2221	.315
3221	.307	1131	.253	1221	.354	3221	.260	1221	.389
2221	.314	2221	.281	2122	.445	2221	.262	1132	.438
1121	.329	3132	.287	1132	.455	1121	.273	2232	.447
1231	.342	1121	.322	2232	.459	1231	.290	2122	.450
3132	.406	1231	.337	2111	.473	3132	.324	2111	.493
1221	.407	3122	.361	1122	.526	1221	.356	1232	.527
3122	.475	3232	.376	2222	.530	3122	.393	1122	.531
2122	.482	3111	.416	1232	.540	2122	.397	2222	.539
3232	.489	1221	.417	1111	.555	3232	.414	1111	.574
2232	.497	2122	.427	2211	.559	2232	.417	2211	.582
3111	.498	2232	.443	1222	.610	1132	.430	1222	.618
2111	.506	3222	.459	1211	.637	3111	.459	1211	.658
1132	.515	2111	.485	2123	.642	2111	.463	1133	.667
3222	.559	1132	.493	1133	.651	3222	.488	2233	.675
2222	.566	3133	.497	2233	.655	2222	.492	2123	.677
3211	.581	3211	.517	1123	.713	1122	.505	2112	.712
1122	.583	2222	.528	2223	.716	1232	.526	1233	.742
2211	.589	1122	.577	2112	.719	3211	.556	1123	.744
1232	.598	3123	.581	1233	.724	2211	.559	2223	.751
3133	.598	2211	.586	1223	.778	1111	.572	1112	.774

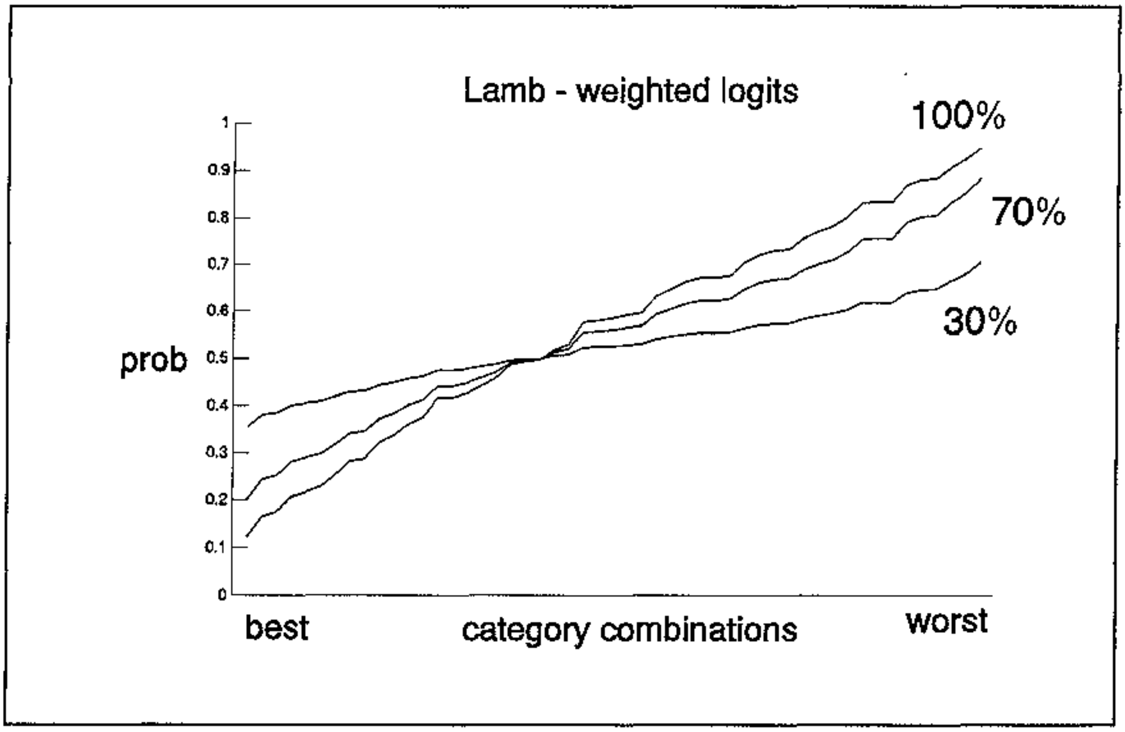
1111	.606	1232	.593	1112	.780	3133	.574	2212	.780
1222	.662	3233	.597	2212	.783	1222	.601	1223	.806
3123	.663	1111	.633	1212	.833	3123	.645	1212	.831
2123	.669	2123	.647	2113	.851	2123	.648	2113	.864
3233	.676	2233	.662	1113	.888	1211	.663	1113	.898
2233	.683	1222	.672	2213	.890	3233	.665	2213	.901
1211	.683	3112	.672	1213	.918	2233	.667	1213	.927
1133	.697	3223	.676			1133	.679		
3223	.734	1133	.704			3112	.698		
2223	.739	1211	.721			2112	.701		
3112	.739	2112	.730			3223	.728		
2112	.745	2223	.733			2223	.731		
1123	.753	3212	.755			1123	.741		
1233	.764	1123	.770			1233	.757		
3212	.799	1233	.782			3212	.773		
2212	.804	2212	.802			2212	.775		
1223	.810	1112	.832			1112	.784		
1112	.815	1223	.834			1223	.808		
1212	.860	3113	.834			1212	.843		
3113	.860	2113	.869			3113	.866		
2113	.864	1212	.881			2113	.868		
3213	.896	3213	.883			3213	.905		
2213	.899	2213	.909			2213	.906		
1113	.905	1113	.924			1113	.911		
1213	.930	1213	.947			1213	.938		

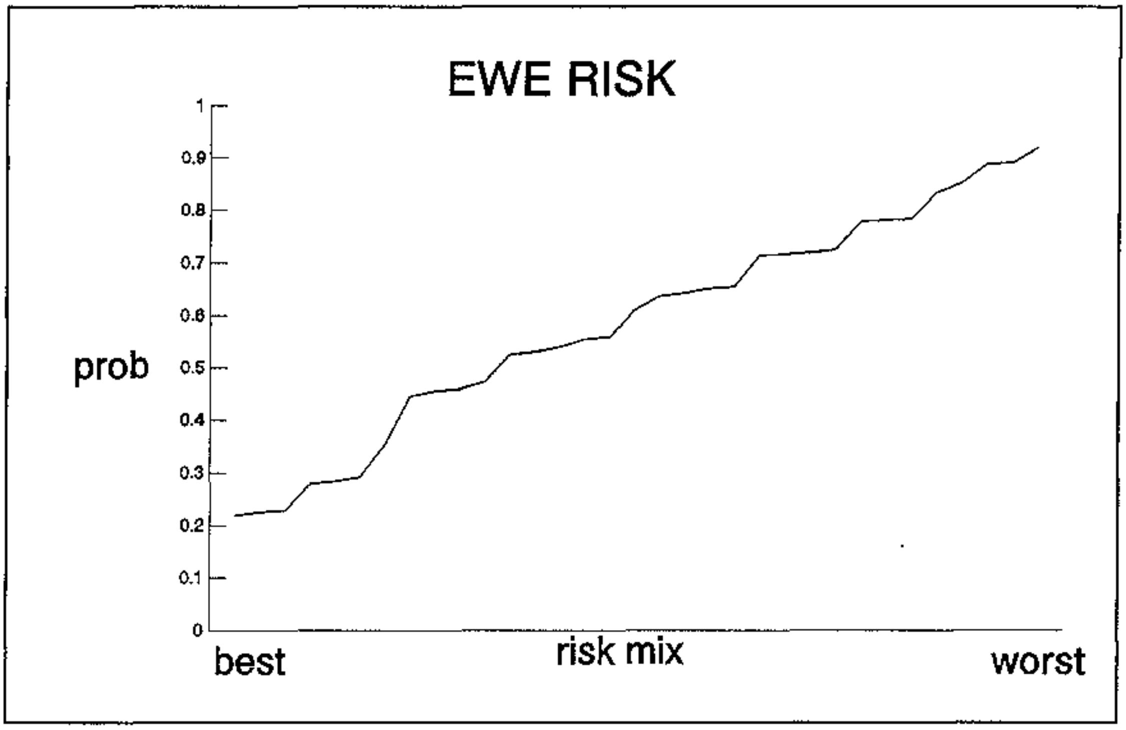
P I A E refers to the price, island, area and expenses category combinations respectively.

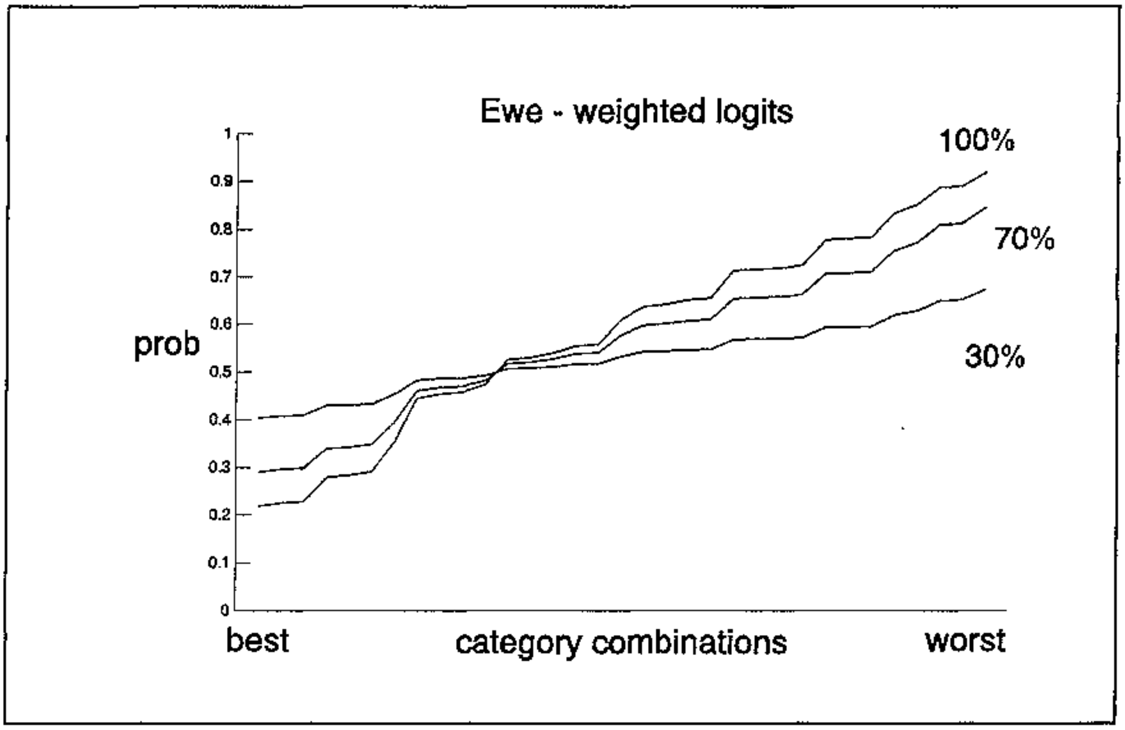
APPENDIX XII

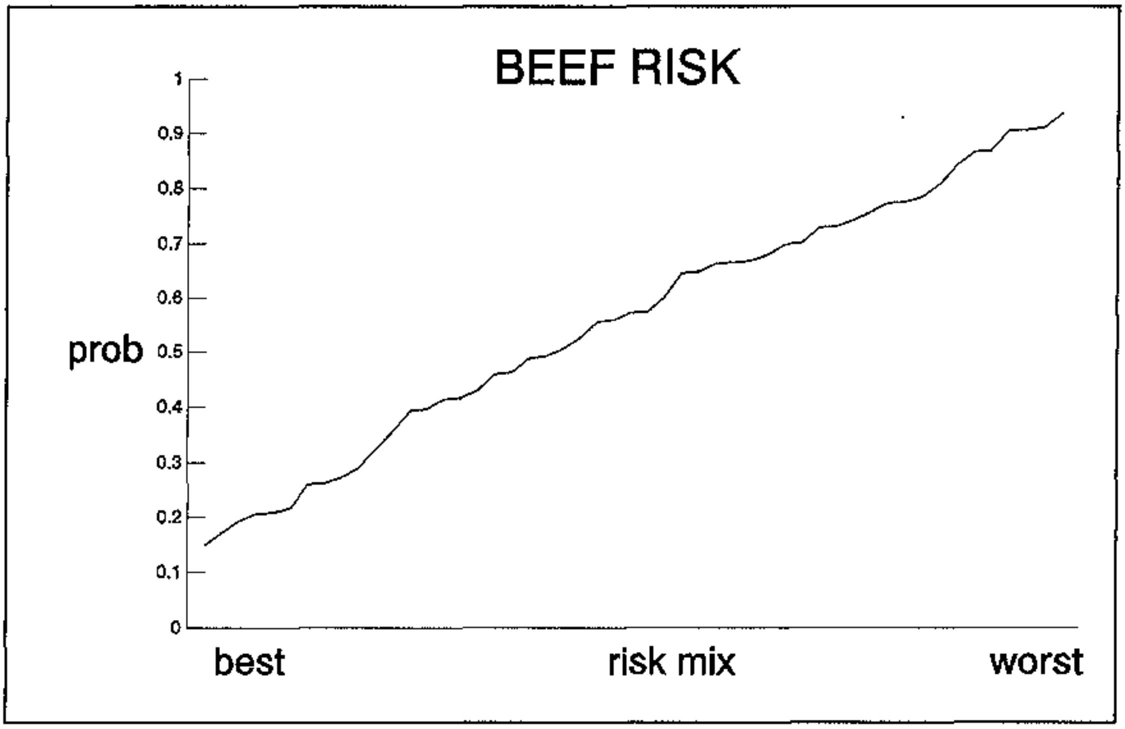
MICRO MODEL RISK FUNCTIONS

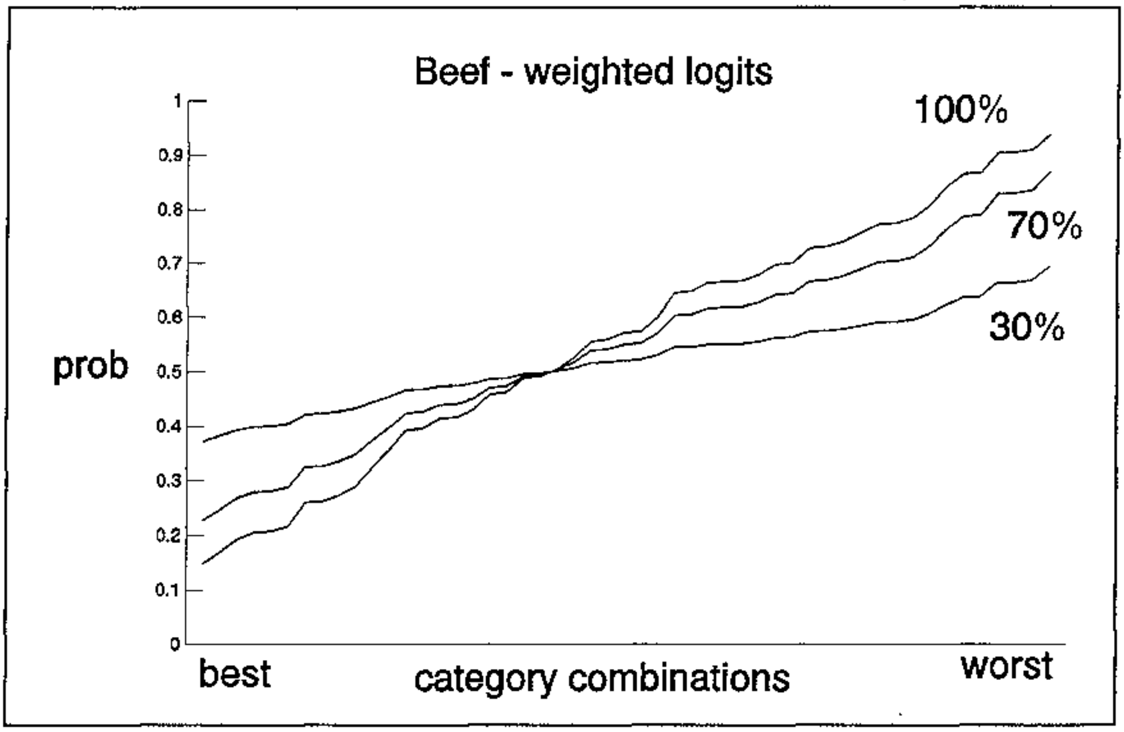


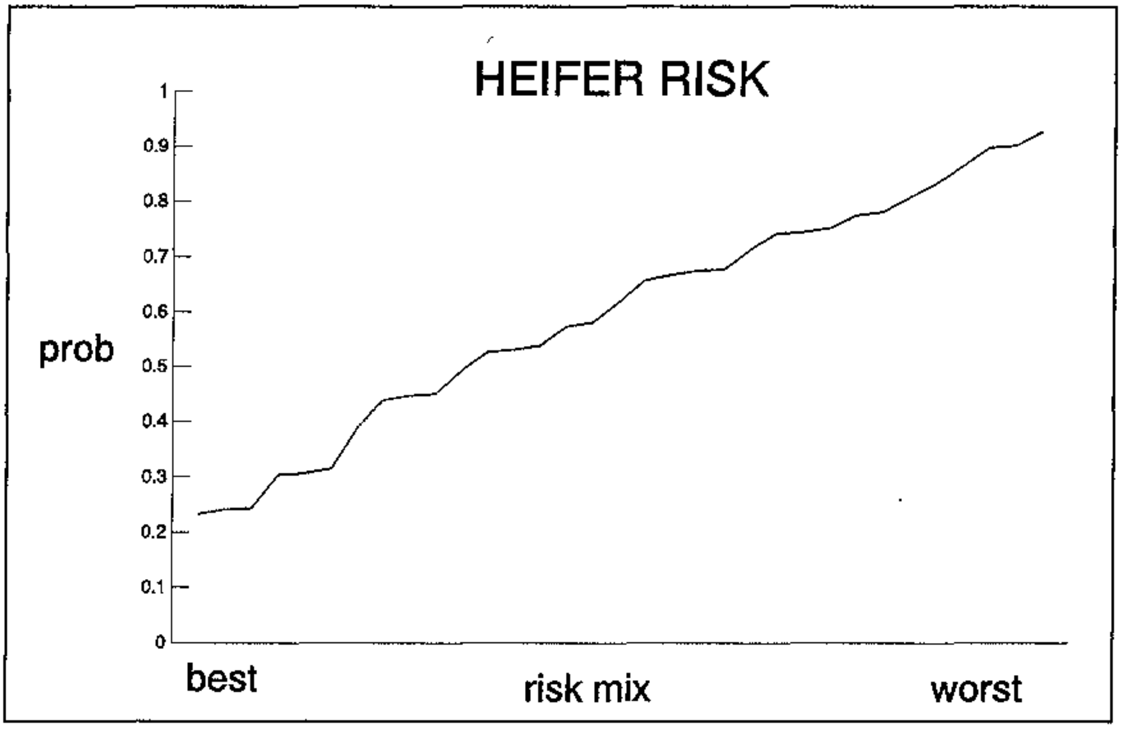


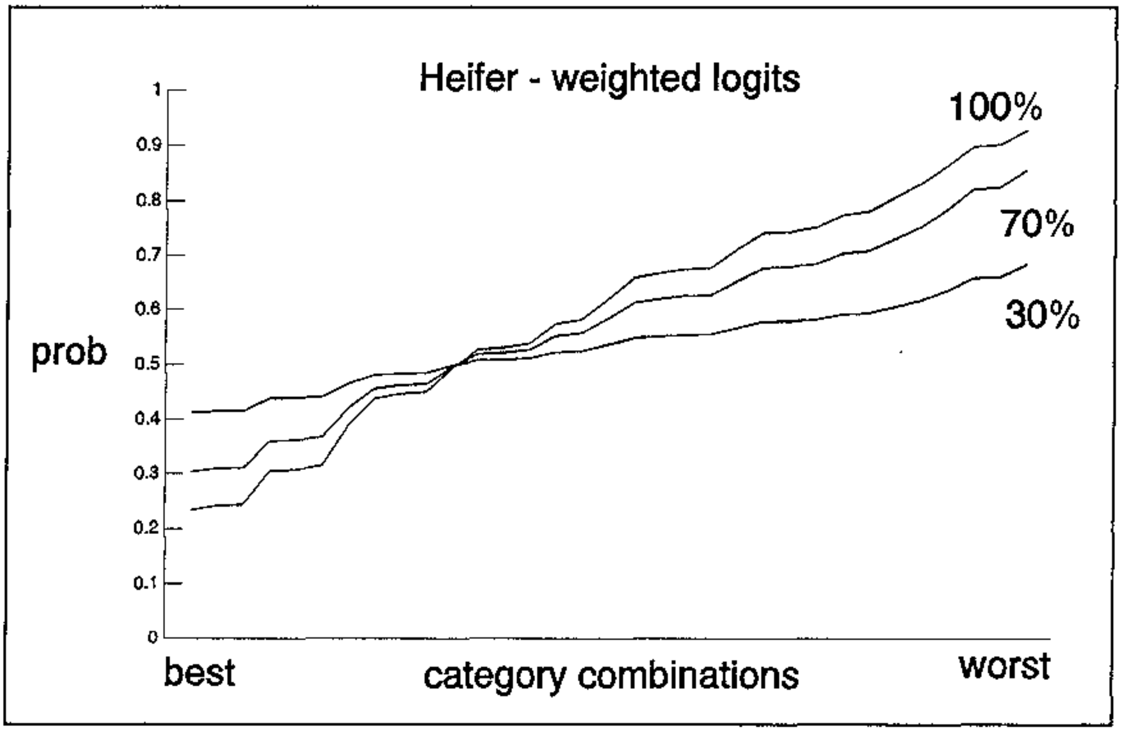

EWE RISK
prob
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
best
risk mix
worst


Ewe - weighted logits
100%
70%
30%
prob
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
best
category combinations
worst


BEEF RISK
prob
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
best
risk mix
worst


Beef - weighted logits
100%
70%
30%
prob
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
best
category combinations
worst


HEIFER RISK
prob
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
best
risk mix
worst


Heifer - weighted logits
100%
70%
30%
prob
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
best
category combinations
worst

APPENDIX XIII.
TEST SAMPLE RISK INDEX DISTRIBUTIONS