COUNTRY RISK:

MULTIVARIATE MODELS AND HUMAN JUDGEMENT

RICHARD ANDREW SOMERVILLE

Thesis submitted for the degree of Ph.D.

The City University

The City University Business School

June 1991

Volume 1

SUMMARY CONTENTS

CONTENTS IN DETAIL

VOLUME 1

VOLUME 2

TABLES: VOLUME 1

TABLES: VOLUME 2

FIGURES: VOLUME 1

FIGURES: VOLUME 2

Acknowledgements

The author wishes to acknowledge the help, patience and encouragement of:

Professor Richard Taffler,
City University Business School;

Mr Fergus Murray,
Computer Laboratory, Trinity College Dublin;

Sue Somerville.

Data was supplied by:

Ms Mina Toksoz, The Economist Intelligence Unit;
Mr Philip Minshull, The Economist Intelligence Unit;
Mr Tom Sealy, International Country Risk Guide.

Declaration

Note on data processing

The data processing for this thesis was carried out on DEC-2060 and VAX-6230 machines at Trinity College, Dublin, Ireland, and also on a Prompt T2 desktop. Most of the statistical analyses utilised the BMDP, IMSL, SHAZAM and SPSS-X software packages on the mainframe installations, while the AID analyses were carried out using the PC-MDS package on the Prompt machine. The specific programmes are referred to where appropriate in the text; the packages are described in Dixon (1988) [for BMDP], IMSL (1989a; 1989b; 1989c), PC-MDS (1988), SPSS X (1986), and White (1982) [for SHAZAM].

Where no reference is made in the text to one of these packages, then processing was carried out using FORTRAN-77 programmes written by the author.

Abstract

The research for this thesis has been conducted at the junction of several intellectual disciplines. These include the theory of human information processing, economics, banking, accounting, and applied statistics, and the results that are reported here should be of theoretical and practical interest to practitioners in all of these fields. This thesis is rooted in the theory of human information processing, and its primary purpose is to test the hypothesis that formal multivariate statistical modelling techniques have a key contribution to make to human decisions. For this purpose, the chosen task-domain is that of creditworthiness assessments by international lenders to less developed countries.

The starting-point for the research is the recent history of lending to less developed countries, and the first original contribution is the derivation of formal multivariate statistical models of credit-worthiness. These are estimated on a set of panel data for a group of 55 LDCs, including all the major debtors. The techniques that are applied include principal components analysis, cluster and proximities analyses, discriminant and logit analyses, and the automatic interaction detector (AID). Using discriminant and logit analysis and AID, early-warning models are derived that perform well, and have a clear operational utility. Moreover, there is a broad consensus across the derived models, concerning variable-specification.

The second major result reported here relates to published and commercially produced ratings of country risk, including those of BERI, the EIU, the ICRG, and the Institutional Investor. The predictive performance of these ratings is shown to be poor. Paramorphic representations of the rating systems are obtained using multiple regression and AID, and these show that the problem is one of bias in the judgemental processes upon which the ratings are based, rather than inconsistency in the application of judgement. This conclusion is supported by an intertemporal statistical analysis of the Institutional Investor rating system.

The performance of the multivariate models is compared with that of the rating systems, and in broad terms the models are found to be superior. In certain circumstances the Institutional Investor system has a lower misclassification cost than the multivariate models, but the findings nevertheless support the case for man-model interaction.

Key to abbreviations

Countries

AL Algeria	HD Honduras	PG Paraguay
AR Argentina	IN India	PE Peru
BD Bangladesh	ID Indonesia	PH Philippines
BL Bolivia	IR Iran	SE Senegal
BR Brazil	IV Ivory Coast	KS South Korea
CM Cameroon	JM Jamaica	SR Sri Lanka
CH Chile	JD Jordan	SU Sudan
CN China	KN Kenya	SY Syria
CG Congo	LB Liberia	TW Taiwan
CB Colombia	MW Malawi	TH Thailand
CR Costa Rica	MA Malaysia	TT Trin. Tobago
DR Dominic. Rep.	MX Mexico	TN Tunisia
EC Ecuador	MC Morocco	TK Turkey
EG Egypt	NC Nicaragua	UR Uruguay
ES El Salvador	NG Nigeria	VZ Venezuela
GA Gabon	PK Pakistan	YU Yugoslavia
GH Ghana	PN Panama	ZA Zaire
GU Guatemala	PP Papua N.G.	ZB Zambia
		ZI Zimbabwe

Variables

Variable labels are defined in Tables IV.2 to IV.5 (pp. 124, 127, 128, 130).

Sundry abbreviations

BERI : Business Environment Risk Information
BIS : Bank for International Settlements
EIU : Economist Intelligence Unit
G7 : Group of seven countries
ICRG : International Country Risk Guide
II : Institutional Investor
IMF : International Monetary Fund
LIBOR: London interbank offered rate
OECD : Organisation for Economic Co-Operation and Development

Introduction and summary

The objectives of this study

The primary issue that this thesis addresses is the performance of multivariate statistical models, relative to that of human judgement, in the context of country-risk assessment, and the thesis makes its original contribution to human knowledge at this junction of several intellectual disciplines. The research includes elements of economics, banking, accounting, statistics and psychology, and its starting point is the recent history of lending to less developed countries. A key aspect of the lending decision of banks and other creditors is the assessment of the creditworthiness of prospective borrowers, and this task-domain is used to test theories of human information processing in an original way.

The results reported here will be of interest to researchers in the several parent intellectual disciplines of this research. Moreover, the thesis has a very practical orientation, and the research findings should be of direct value to lenders and their regulators. This is true of the general conclusions concerning human information processing.

Equally, although the thesis is not primarily about early-warning models, the derived models should nevertheless be of interest and use to bankers, bank regulators, and official lending agencies.

Accurate risk assessment is clearly in the interest of banks and their shareholders, and of official agencies and taxpayers in the industrial countries. It is equally in the interest of prospective borrowers. If lenders are unable to distinguish between high- and low-risk countries, then countries in the former category may become burdened with debt that they will be unable to service. Alternatively, low-risk countries may find themselves denied flows of capital that they could use efficiently and profitably.

The objective of this study is to test the hypothesis that a multivariate statistical approach using formal modelling techniques can be a useful tool for country risk assessment. The benchmark for assessing the multivariate models is human judgement, as exemplified by the Institutional Investor country credit ratings, the International Country Risk Guide (ICRG) risk ratings, the Economist Intelligence Unit's credit risk rating scores, and the Business Environment Risk Information (BERI) index.

Literature surveys

The ancestry of this thesis in several distinct disciplines is reflected in the literature that will be cited in later chapters. Much of this is surveyed in chapter II, of which a major focus is the theory of human information processing, in which this thesis is rooted. However, that chapter also surveys the application of multivariate statistical forecasting models to a number of task-domains, including sovereign lending risk, political risk and corporate risk. In later chapters applications of various statistical techniques are reported, and the statistical literature is referenced as appropriate.

Structure of thesis

The chapters in which the arguments are developed fall into three groups. Chapters I to V present the framework of the thesis: they include a description of the LDC debt problem, a review of recent research on human information processing, and the data. In the next group of chapters, VI to XI, multivariate models are developed and tested. Finally, chapters XII and XIII deal with expert judgement and its paramorphic representation in the case of country risk assessment.

The two concluding chapters draw the earlier arguments together: Chapter XIV compares the classificatory performance of multivariate models with that of expert judgement, and finally the thesis is summarised and conclusions are presented in chapter XV. The chapters are described more fully in the following sections.

The framework

Chapter I summarises the recent history of LDC debt, in order to put the rest of the thesis in context.

The modelling approach is rooted in the theory of human information processing, and this theory, which is a novel approach to country risk assessment, is surveyed in chapter II. The proposed hypothesis is that formal multivariate statistical modelling techniques have a key contribution to make to human decision processes. Man is prone to judgemental biases, and these may be reduced or eliminated through a formal modelling approach.

Chapter III is concerned with the event to be predicted. This is taken to be the onset of arrears, whereas earlier work in this area utilised the event of rescheduling or restructuring. This novel

approach is intended to generate models that will be of more practical utility than their predecessors.

Chapters IV and V describe the data: its sources, its composition, and its distributional properties. Compared with the data used in earlier studies, it includes a wider range of variables. Insight into the dimensions of the data is provided by principal components and cluster analyses.

Multivariate models

Chapters VI to XI describe the specification, estimation and testing of multivariate models. The statistical approach of these chapters, and also of chapters XII and XIII, is based on the 'general to simple' methodology that is discussed in Gilbert (1986) and Hendry and Mizon (1985), who write (p.3):

> Science is a process of accumulating empirical knowledge based on a data-instigated progressive research strategy in which the model is the message.

Broad initial specifications are assumed; these are then narrowed down using stepwise procedures, and at every step the specification is tested for congruence with the data. Holdout samples are retained for out-of-sample validation of the derived models.

The statistical methodologies employed include cluster and proximities analyses, discriminant analysis, logit analysis, and automatic interaction detection (AID). Most of the earlier research in this area has used either logit or discriminant analysis. In this thesis, both techniques are used, the maintained distributional assumptions for the discriminant model are exhaustively explored, and the derived models are compared and found to be very similar. AID has not been used in this context before: it does not replace the other approaches, but provides key information about interaction effects, and confirmatory evidence for the specifications of the other models.

Using three different approaches to testing, the models are found not to be sample-specific. Moreover, the classificatory performance of the derived models is found to be satisfactory, through validation on a holdout sample.

Human judgement

Chapters XII and XIII are concerned with banker and expert judgement, as reflected in commercially produced and published ratings of country risk. Chapter XII explicates banker judgement in terms of

underlying economic variables. The technique has been applied elsewhere to the analysis of bond ratings, but this is the first application to ratings of country risk, using regression and AID analyses. It is therefore an original contribution of this thesis. It emerges that the different rating systems embody different judgemental bases, and this point is explored further using a correlation analysis. Moreover, the judgemental bases of the ratings involve information sets that differ from those specified for the multivariate models.

Chapter XII includes two other findings concerning expert judgement. First, an intertemporal analysis of the Institutional Investor rating suggests that bankers may be biased towards inadequate revision of their opinions in the face of new evidence: thus, the 'anchoring and adjustment' heuristic may be at work. Secondly, evidence is presented that perceptions of political and economic risk are not perfectly correlated.

Chapter XIII compares predictions yielded by ratings ('expert judgement') with those yielded by paramorphic representations of expert judgement. This is a novel setting for the application of a technique that has been applied to other tasks, as a means of detecting and correcting for random error

and bias.

The paramorphic representations do not improve on the ratings upon which they are based. In the case of those ratings for which reasonably large sets of data are available, the classificatory performance of the ratings and their paramorphic representations is about the same, and is poor in either case. This is a key result: the issue is bias, not lack of consistency. Poor judgement is being applied consistently, and this emphasises the need for improved techniques for assessing country risk.

Concluding chapters

Chapter XIV draws together the material on multivariate models and human judgement. Within the training sample period 1979 to 1986, the multivariate models have a superior predictive performance, compared with expert judgement or its paramorphic representation. On the holdout sample, the position is more complex. Under certain circumstances, banker judgement (represented by the Institutional Investor rating system) has a lower misclassification cost, but it is seen to be a rather blunt instrument: type I errors are eliminated at the cost of a very high incidence of type II errors. This supports the case

for using multivariate models in conjunction with human judgement: they should be seen as complementary, not competitive.

The major conclusions of chapter XIV are presented in chapter XV, along with subsidiary conclusions from other chapters, and a consideration of the audience for these research findings. The main conclusions of the thesis may be summarised briefly here.

Ratings of expert and banker judgement, which have distinct elements of **economic risk** and **political risk**, are biased but are applied consistently. Simple multivariate statistical models achieve an impressive performance, which in broad terms dominates that of the rating systems. Thus the multivariate models have a direct practical utility, although there is also a role for man-model interaction. Finally, the bootstrapped models of the rating systems suggest a judgemental basis that is distinct from the specification of multivariate early-warning models.

I Background to the sovereign debt crisis

I.1 Introduction

The problem that this thesis addresses is the assessment of the creditworthiness of less developed countries (LDCs). In recent years, most of these countries have experienced difficulties in servicing their external debts, and this chapter will survey these events.

Section I.2 outlines the events that preceded the breaking of the crisis in 1982, while I.3 and I.4 summarise events since then. Section I.5 deals with inter-regional differences in debt-servicing problems. Section I.6 is concerned with the impact of the crisis on creditor banks, and a brief summary of the chapter is presented in section I.7.

I.2 Events before 1982

There is nothing new, or necessarily undesirable, in the amassing of external debts by developing countries. Many of the industrial countries relied heavily on capital inflows during earlier phases of their own development, to finance levels of fixed investment that exceeded domestic savings. However, since 1982 there have been widespread difficulties in the relationships between debtor and creditor countries. The origins of these difficulties can be traced back to 1973, and it is appropriate to begin this study with a résumé of events since that year.

Until 1973, the external debts of less developed countries (LDCs) consisted largely of long-term loans. Many were project-related, and issued on concessional terms, and most were issued by official agencies: for example, governments in the OECD countries, or multilateral institutions.

After 1973, commercial banks in the OECD countries became a progressively larger source of credit to the LDCs, up to about the middle of 1982, as shown in Table I.1.

Table I.1: Developing countries: composition of external debt and borrowing: 1973-88

	1973	1979	1981	1982	1983	1986	1988
External debt $bn	130	391	752	852	904	1114	1240
Of which:							
(a) Guaranteed long term debt to banks per cent:	--	--	25	25	28	31	28
(b) Short term (1) + other private (2) per cent:	--	--	46	45	41	34	32
(a) + (b)	61	66	71	70	69	65	60
Net borrowing $bn	11	48	126	104	77	47	26
Of which:							
Commercial banks per cent:	[47]	[69]	78	62	42	-8	-27

Sources:
IMF, 1981: Table 22
IMF, 1983: Table 1
IMF, 1989b: Tables A40,A46

Notes:
(1) Includes official short term debt.

(2) Includes unguaranteed long term debt to official creditors (if any). Most unguaranteed long-term debt is owed to private creditors, including banks. Category includes guaranteed long term debt to nonbank private creditors.

[ ] includes all private sector creditors and all short term debt.

-- Data not available.

The maturity of the outstanding debt fell, and an increasing share carried a floating and market-related rate of interest. The outstanding debt, in aggregate and also for most individual debtor countries, grew enormously.

The changes that occurred after 1973 arose in the first instance from the two oil 'shocks' of 1973-4 and 1979. These shocks had at least two major implications for the non-oil LDCs. First, they adversely affected their terms of trade. Secondly, they had a major recessionary impact on the OECD countries, which accounted for about 75 per cent of world product and were the destination of about 70 per cent of the exports of the LDCs. When the OECD area moved into recession, the demand for most exports of the LDCs fell, with a particularly severe impact on the prices of primary commodities other than oil.

These events caused severe difficulties for the balance of payments of the non-oil LDCs. At the same time, the oil exporters were accumulating external surpluses, which were largely deposited with commercial banks in the OECD area. The process whereby these were lent to the non-oil countries to finance their current account deficits became popularly known as the 're-cycling of petro dollars'.

The LDCs' balance of payments positions were also adversely affected during the 1970s and early 1980s by the growth of protectionist policies within the OECD area. These included tariffs and quotas, and certain types of explicit and implicit subsidies on exports.

In 1980, the forces that were tending to increase the debts of the LDCs were augmented by rising real interest rates, which resulted from tight monetary and expansionary fiscal policies in the United States. This increased the burden of debt servicing, and provided an impetus for further borrowing.

The extent of the difficulties faced by any given country depended, inter alia, on its economic policies, including its reactions to external shocks. Some countries that suffered worsening terms of trade in the 1970s made appropriate adjustments and did not become problem borrowers. Examples in this category are South Korea and Thailand. Conversely, debt problems developed in other countries, such as Mexico and Brazil, where terms of trade effects were weak or even favourable.

A common inappropriate policy response consisted of attempts to maintain public consumption at pre-shock levels on the basis of large fiscal deficits financed

through external borrowing. Secondly, many countries attempted to maintain overvalued exchange rates, thus eroding the competitiveness of domestically produced traded goods. In general, the immediate result of overvaluation is deterioration in the balance of payments, while the longer-term effect may be an undermining of the tradeable-goods sector of the economy.

A further effect of overvaluation is the promotion of capital flight. For example, according to estimates made by Morgan Guaranty Trust Company, aggregate flows of flight capital in 1976-1985 for the major Latin American debtors were: Argentina $26bn, Brazil $10bn, and Mexico $53bn. (IMF, 1986a: table 52.) Where, as commonly occurred, capital flight accompanied a rising debt, then part of the new debt was in effect incurred through the purchase of external assets by private residents, with no prospect that any of the income thereby generated would be available for debt servicing. Finally, trade regimes in the countries with debt problems were typically biased against exports, particularly those of agricultural products.

I.3 Developments since 1982

During the 1970s debt was growing, but for most LDCs, the debt/export and debt/GDP ratios were fairly stable. However, the aggregate data that are set out in Table I.2 show that the burden of debt and debt service started to climb in about 1981, and that this continued through to 1986.

Table I.2: Less developed countries: ratios of total external debt service to exports and GDP

	1973	1979	1981	1982	1983	1986	1988
				Per cent			
debt/exports	115	117	96	120	135	171	142
debt/GDP	22	27	28	31	33	38	36
debt service/exports	16	19	16	20	19	23	20
debt service/GDP	3	4	5	5	5	5	5

Sources:
IMF, 1981: Table 1
IMF, 1989b: Tables A48,A49

From 1979, a growing number of countries experienced difficulties in servicing their external debts, and in August 1982 the Mexican government declared a moratorium on amortisation payments to creditor banks. According to Cline (1983), 34 countries were in arrears by the end of 1982.

In the aftermath of 1982, the reactions of creditors and debtors reflected a view that the problem was essentially a temporary liquidity crisis, connected with a mis-match between the timing of debt-service commitments on the one hand, and of expected returns from productive investment in fixed assets on the other. Thus, debtors attempted to cut their current account deficits through cuts in imports, and allowed arrears of debt service to develop. In reaction to arrears, and in anticipation, creditors and debtors entered into rescheduling arrangements.

Since the commercial attractiveness of lending to LDCs was much reduced, the banks cut back on new lending. Net external borrowing by the LDCs from all sources fell from $126bn in 1981 to $77bn in 1983, from which point it fell to $26bn in 1988. The banks' share of this total collapsed from 78 per cent in 1981 to the point where net repayments were made to the banks by the LDCs in 1986. In the latter year, the banks' share of net lending was -8 per cent. After a re-emergence of net bank lending in 1987, the banks' share was again negative, at -27 per cent, in 1988 (see Table I.1).

As the 1980s proceeded, it became clear that the debtors and creditors were not faced merely with a temporary crisis of liquidity.

Table I.3: Secondary market prices: LDC loans

	June 1986	Sept 1987	Sept 1988
Prices: US cents per $ of nominal value:			
Argentina	64.0	37.0	21.8
Bolivia	6.0	9.0	10.0
Brazil	75.0	39.0	46.3
Chile	66.0	58.0	59.5
Ecuador	63.5	33.0	21.0
Ivory Coast	74.0	60.0	26.5
Mexico	60.0	47.3	46.8
Nigeria	55.0	25.0	27.0
Philippines	59.0	59.0	52.5
Turkey	97.0	96.5	98.5
Venezuela	76.0	63.0	51.0
Yugoslavia	79.0	60.0	47.0

Sources:
World Bank, 1988-89: Table 7
Saloman Brothers Inc., 1987

The perception grew that more fundamental structural adjustment was required within the debtor economies. This perception has been reflected for some years now in the heavy discounts at which LDC debt is traded in the secondary markets, as illustrated in Table I.3.

I.4 Reactions to the crisis

The problem of external debt has two principal dimensions: first, the domestic economic implications of servicing high and rising levels of debt, and secondly the problems that arise for debtors and creditors when debt service payments fall into arrears.

Sovereign debt consists of debt incurred by the state and its agencies, and the LDC debt problem is sometimes regarded as a problem involving sovereign debt only. However, private debt also generates a demand for foreign exchange to meet debt servicing commitments. According to recent IMF estimates (IMF, 1989b: Table A46), private unguaranteed debt declined from 23 per cent of the total outstanding long term debt of LDCs in 1981 to 12 per cent in 1988. A significant factor in this development has been the ability of creditor banks, during the course of restructuring negotiations, to secure the extension of guarantees from the governments of LDCs to cover the debts of private debtors. Chile has been an outstanding example of this phenomenon: as Griffith-Jones (1989: 60-61) has written, the Chilean government has become, ex post, 'borrower of last resort' relative to the debts of the Chilean

financial sector. Private unguaranteed debt accounted for 64 per cent of long term external Chilean debt in 1981, and only 14 per cent in 1988 (World Bank, 1988-89).

A country is unlikely to fall into debt problems if its economy is basically strong and diversified and if its external borrowing respects three principles. First, debt should be used only to finance assets whose rate of return exceeds their financing costs. Floating interest rates have made this a difficult target in recent years. Secondly, the maturity of debt should not be grossly mis-matched with the timing of receipts from the corresponding asset. Thirdly, the foreign exchange composition of total debt service should broadly match that of foreign exchange earnings.

In contrast with the first of these principles, many countries responded to adverse changes in their trading environment by accumulating debts, in order to maintain current domestic spending at levels that would otherwise have been unsustainable. At the same time, supply-side stimuli to domestic tradeable goods sectors were inadequate or even negative.

Table I.4 shows how the total debt of the LDCs has changed since 1981. Between 1981 and 1982, it grew

by 13 per cent. In the aftermath of 1982 the rate of growth moderated, but it never fell below 5 per cent per annum until 1988, when the year-on-year growth rate was 0.7 per cent.

From 1981 to 1986, the total external debt of all LDCs grew faster than either exports or GDP: this is reflected in the rise in the ratios of debt to exports and GDP, as set out in Table I.2, where it may also be seen that some improvement has taken place since 1986. A similar pattern is revealed by the behaviour of the debt service ratio.

Table I.4: LDCs: growth rates of total external debt and debt service

	1982	1983	1984	1985	1986	1987	1988
	Annual average rate, per cent.						
Total debt	13.2	6.2	4.7	8.1	8.9	10.5	0.7
Total debt service	10.3	-10.8	12.6	2.8	3.7	5.0	8.7

Source
IMF, 1989b: Tables A46, A49

Thus even after 1982, it was certainly the case until 1986 that the burden of debt of all LDCs taken together was growing - both in absolute terms, and relative to the ability of the LDCs to service it. The same was true of the debt service commitments of the debtors, as distinct from actual payments of debt

service. A natural reaction of creditor institutions in such circumstances was to revise their assessments of the creditworthiness of borrower countries. The supply of new lending contracted, and lenders became less willing to roll-over maturing debt. They also tended to seek shorter maturities and higher margins on such new lending as they were still prepared to offer. The natural conclusion of such a sequence of events is a liquidity crisis, and the events of August 1982 were the first example of this happening on a major scale in recent history.

Shortages of liquidity have been the experience of many LDCs since the oil 'shock' of 1979. During the early 1980s, the external debt of many LDCs was so high that the cost of its servicing was a major constraint on economic progress. Typically, highly indebted LDCs were faced with the necessity of undertaking large-scale structural adjustments on both the supply and demand sides of their economies. The required adjustments would include significant increases in the resources devoted to export-orientated investment. The cost of this, in terms of reduced domestic consumption, could be moderated to the extent that the new investment was externally financed. However, the longer the delay in implementing the required adjustment, the greater would be the eventual adjustment required, the

greater would be the magnitude of the accumulated debt burden, and the lower would be the willingness of creditors to underwrite even worthwhile investment projects. Such is the predicament of many LDCs at the start of the 1990s.

The existence of a large external debt may in itself constitute a disincentive to domestic adjustment. The actual debt service payments of heavily indebted countries tend in practice to be related to economic performance. This implies that the benefits from improved domestic policies will largely accrue to external creditors, and this effect, known as 'debt overhang', may make it difficult or impossible for governments to obtain political support for policy changes that would increase output and exports. A recent IMF study (IMF, 1989b: supplementary note 1) concludes that this hypothesis cannot be dismissed on empirical grounds.

I.5 Regional patterns

The IMF uses five geographical regions to classify less developed countries. These, together with their share of total LDC debt outstanding in 1988, are:

Western Hemisphere	33 per cent
Asia	26 per cent
Africa	16 per cent
Europe	13 per cent
Middle East	12 per cent

This ranking, although not the size of the shares, was the same as in 1981. The ranking was also the same in the intervening years, except that the Middle East and Europe were reversed in order during 1983-85.

One-third of LDC debt is concentrated in the Western Hemisphere - i.e. central and south America and the Caribbean. Moreover, the three largest debtors, Brazil, Mexico and Argentina, have been the three biggest problem cases for creditor institutions taken globally. Other major problem cases in this region are Venezuela and, at least until recently, Chile, and of the World Bank's category 'fifteen heavily indebted countries' (the 'Baker fifteen'), ten are in Latin America.

The second ranking region is Asia. However, Asia contains a number of strong economies that have built

up diversified manufacturing and primary sectors. South Korea is the fourth largest debtor among all LDCs, but it has never been a problem for its creditors. The Philippines is the only Asian member of the Baker fifteen, and it has been the only Asian problem case in recent years.

Africa includes three of the Baker fifteen, namely Ivory Coast, Morocco, and Nigeria, together with such other problem cases as Zaire and Zambia. For reasons of scale it tends to receive less attention than the Western Hemisphere. However, in terms of the burden of debt on the debtor countries, Africa, particularly south of the Sahara, includes some very hard cases. It includes many low income countries, and a low income per head sets a severe limit to the capacity of a country to withstand adverse trading conditions. Many African countries have relatively simple economies that are in many cases reliant on a very few export commodities: such countries are particularly exposed to external shocks. For example, copper accounted for 88 per cent of Zambia's exports in 1984-85. At the same time, iron ore accounted for 63 per cent of the exports of Liberia, while 66 per cent of Ghana's exports consisted of cocoa beans (United Nations, 1988: Table 4.3).

Of the 55 countries that are included in the

statistical analyses that are reported below, 49 are drawn from the first three regions and only six from Europe and the Middle East. Taking the last two regions together, the major problems are in Europe - namely, Poland and Yugoslavia. Romania, formerly a problem case, has recently become the first recent example of such a country paying off virtually all of its external creditors. The only member of the Baker fifteen in these two regions is Yugoslavia.

Table I.5: LDCs: ratios of total external debt to GDP and exports; regional analysis

	1981	1982	1983	1984	1985	1986	1987	1988
Debt/exports:				Per cent				
Western Hem.	210	272	293	275	297	353	346	305
Asia	74	87	93	88	102	103	89	76
Africa	119	155	171	171	191	245	249	249
Europe	137	145	149	145	161	168	169	146
Middle East	35	48	64	72	85	116	110	117
All LDCs	96	120	135	134	151	171	159	142
Debt/GDP:				Per cent				
Western Hem.	40	44	47	47	45	46	46	44
Asia	19	21	23	23	26	29	28	25
Africa	31	35	38	42	47	50	49	50
Europe	34	35	37	41	45	43	45	42
Middle East	19	23	24	26	27	31	31	31
All LDCs	28	31	33	34	36	38	38	36

Source
IMF, 1989b: Table A48

Table I.5 sets out debt-ratios for all five regions for the years 1981 to 1988. The overall picture is

one in which the ratios of debt to exports and to GDP rose up to 1986, and fell thereafter. Apart from Africa, the regions followed this pattern. The greatest deterioration up to 1986 occurred in the Middle East, and the greatest improvement since then has occurred in Asia. In Africa, the ratio debt/GDP fell slightly in 1986-87, and then returned to its 1986 level in 1988, while the ratio debt/exports continued to rise up to 1987, then fell by 0.1 per cent in 1988 to a level that still exceeded that reached in 1986.

I.6 The banks' problem

The growth of bank debt is central to the debt crisis. As a group, banks in the industrialised countries have been accused of causing the debt crisis by engaging in unwise lending practices. In the banks' defence, it must be acknowledged that country lending was a new business for them in the 1970s: in earlier periods, external debt consisted largely of bonds. Furthermore, in the 1970s the OECD governments and the international organisations were in favour of the 'recycling' of petroleum surpluses. As problems developed, the larger and most heavily exposed banks faced a dilemma: any single bank that tried to leave the market could have precipitated a crisis, while the alternative was to continue lending in decreasingly attractive circumstances.

The problem of bank debt is largely centred on the middle income developing countries, particularly in Latin America. Although the exposure of BIS-reporting banks in Latin America has been falling in recent years, their total loans outstanding to LDCs in the Western Hemisphere in June 1988 still accounted for 47 per cent of all loans to borrowers in the capital importing developing countries. Most (88 per cent) of this exposure in the Western

Hemisphere was accounted for by Mexico, Brazil, Argentina, Venezuela and Chile, in that order, so that these five countries alone accounted for 41 per cent of outstanding bank loans to capital importing developing countries. By contrast, Africa south of the Sahara accounted for below 7.5 per cent of the total. (IMF, 1989a: Table A29).

Asia has been the only region to which the banks have been net lenders in every year since 1981. Africa and the Middle East started to make net repayments to the banks in 1984, other regions followed, and overall there were net repayments in 1986 and 1988. Notably, whereas net lending in the Western Hemisphere was $60bn in 1981 and $38bn in 1982, net repayments from that region were $12bn in 1988.

Bank credit commitments tell a similar story. Commitments to all LDCs fell from $42bn in 1982 to $19bn in 1988. Of this reduction of $23bn, 70 per cent was borne by LDCs in the Western Hemisphere.

According to Cline (1983,p.36), writing in the aftermath of the Mexican moratorium of 1982:

> For Western banks, repudiation of a substantial portion of loans to developing countries and Eastern Europe would be crippling. Even widespread moratoria could have a severe impact on the banks.

Since then, the exposure of the banks to developing

countries has declined. For example, external claims of all US banks on developing countries peaked in 1983, and have since declined steadily. At the same time, total assets, and capital, continued to grow. Thus, the ratio **capital/external claims on developing countries** reached a trough in 1981-82, and has since grown steadily from 47 per cent then to 138 per cent in 1988. The share in total assets of external claims on developing countries has correspondingly fallen since 1982. (IMF, 1989a: Table 7).

The strength of the aggregate balance sheet of the US banking system, as measured by the ratio of capital to total assets, has risen continuously from a low point of 5.3 per cent in 1980 to 8.1 per cent in 1988, and this improvement has been matched in other major industrial countries including Canada, France (since 1985), Germany, Netherlands, Switzerland, and the UK, but excluding Japan (IMF, 1989a: Table 9).

The declining exposure to developing countries has come about via two routes, apart from the impact of exchange rate changes. First, the banks have reduced their new loan commitments, and in consequence net lending has been negative in some recent years (see above). Secondly, the banks have used various means to reduce the LDC loans on their books. Since 1982, a secondary market has emerged, and some banks have

disposed of loans by this means, albeit at a discount. Other similar techniques have included direct buy-backs at a discount, debt-equity swaps, and the swapping of debt for fixed- interest bonds. All these measures imply a capital loss for the banks, but have the attraction that they remove the banks' liability to provide new loans. According to a recent estimate (The Economist, 18 March 1989: p. 110), the international banks reduced their stock of Latin American debt by about $17bn in 1988. Bolivia and Chile have been particularly prominent in reducing their indebtedness by these means.

Alongside these techniques of debt-extinction, many of the major creditor banks have made balance-sheet provisions against loans to developing countries. Citicorp was the first to make a large specific provision against against LDC loans, when it announced a provision of $3bn in May 1987 (Citicorp, 1987). This was quickly followed by similar action by the other major international banks. Recently a British all-party parliamentary group concluded:

> American banks are now more vulnerable to domestic energy, farming and housing loans than to LDC debt. For most of the major banks, simultaneous default (collective or coincidental) by a number of large Latin American debtors would shake them; a single default would be absorbed. (Griffith-Jones, 1989: p.43.)

I.7 Conclusion

This chapter has outlined the recent history of the LDC debt crisis, including its implications for creditors and debtors. The total stock of debt has continued to grow in recent years, although the burden (as reflected in debt/export and debt/GDP ratios) has been falling since 1986. Strengthening balance sheets leave the creditor banks somewhat less exposed than hitherto to the risks of default or repudiation.

II Human information processing: models and man

II.1 Introduction

The primary concern of this chapter is with the efficient use of information in human decision processes, with particular reference to assessments of credit risk. Sections II.2 and II.3 survey recent research on human information processing. The conclusions of this research supply a rationale for the use of quantitative techniques, and this will be explored in sections II.4 and II.5. Section II.6 surveys recent quantitative approaches to the assessment of country and corporate risk, and the chapter concludes (section II.7) with a summary of the evidence on man-model interaction.

II.2 Human information processing

II.2.1 Relation to objective of thesis

The external obligations of private and official debtors in developing countries are owed to banks and other private creditors, governments, and multilateral agencies such as the World Bank. Creditors have a natural, and vital, interest in country risk analysis, and commonly undertake it internally. Central banks, universities and other research bodies are also actively involved. Finally, the financial press has a continuing involvement, notably Euromoney and Institutional Investor. Thus, numerous researchers in diverse institutions throughout the world are engaged in assessing the creditworthiness of debtor countries.

This research may be classified, following Goodman (1977), as follows:

(i) fully qualitative

(ii) structured qualitative

(iii) weighted checklist

(iv) other quantitative

The first category includes methods that are based on a qualitative country report, of unstandardised format and structure, that is evaluated subjectively.

The second is a modification of this in which the format is standardised and in which there is some summary statistical analysis. In the checklist system, a set of variables is scored and aggregated into a summary rating, using subjectively determined weights. The final category includes formal multivariate and econometric techniques, among which are 'early warning' models of the sort that are described in later chapters. All these categories involve the use of statistical information as an input to the process of assessment and decision-making, but of course they differ in the manner in which the information is processed.

A central concern of this thesis is the use of information. The remainder of section II.2 will survey the theoretical literature on human information processing, while relevant empirical findings will be reported in section II.3.

II.2.2 The Brunswik lens model

Brunswik's lens model (Brunswik, 1952) is concerned with a decision maker or analyst who is separated in time and space from the event that he is analysing. He receives multiple and overlapping cues that are imperfect predictors of that event, and he must

determine how to combine and read the cues so as to obtain a probabilistic estimate of the event. In Brunswik's conceptualisation, these cues, which are probabilistic, form a lens that separates the analyst from the task environment in which the event resides. The accuracy of the analyst's judgement depends jointly on the structure of the environment and on his own characteristics.

All the approaches to country risk analysis that are listed in II.2.1 above may be cast into Brunswik's lens framework. Evaluation and comparison of them requires us to focus jointly on the relationship between the event in question (e.g. restructuring of debt) and the cues, and also on the manner in which the analyst utilises the cues.

II.2.3 Heuristics and biases: introduction

The categories in Goodman's classification (see II.2.1) are shown in order of increasing formality. Using fully qualitative techniques, the analyst is working as an 'intuitive statistician'(Peterson and Beach, 1967). The cues may or may not include statistical data in organised format, but the essential features of these analyses are that they do not follow a standardised plan, and that the

judgements are reached in a subjective manner.

A large body of research in human information processing is concerned with the hypothesis that 'man as an intuitive statistician' is liable to incur significant biases in his use of information: see, for example, the collection of papers in Kahneman, Slovic and Tversky (1982). Tversky and Kahneman (1974) describe three heuristics (i.e. 'rules of thumb') that they claim to be commonly used in the making of judgements about the probabilities of occurrence of uncertain events. Each is a potential source of bias, in that each of these heuristics may be less than perfectly correlated with the variables that determine the events in question.

II.2.4 Anchoring and adjustment

The anchoring and adjustment heuristic is the practice of forming estimates by making an adjustment to some initial value or position. Different initial values yield different estimates, and according to Tversky and Kahneman the tendency is for estimates to be biased towards initial values. This arises from conservatism, manifested as a failure to make sufficient adjustment to initial values, when new evidence should stimulate a more radical

reassessment.

Evidence exists that this heuristic may guide bankers in their country analyses. For example, the second oil shock of 1979-80 had a significant impact on debt-servicing capacity, but the outturn varied according to individual countries' specific economic characteristics and policies. Despite the new information conveyed by the oil shock, bankers' views on relative credit-worthiness, as revealed by the Institutional Investor rating, appear to have changed very slowly. This evidence is presented in chapter XII.

A second example of anchoring and adjustment is the common practice of forming forecasts of economic variables by extrapolating from recent or current values. This procedure generally misses turning-points, and embodies a serious bias towards initial values.

II.2.5 Availability

The availability heuristic is employed when, in forming a view about the future course of events, decision takers are influenced by the ease with which the possible outcomes can be envisaged. Tversky and

Kahneman (1974: p.11) point out that availability is

> a useful cue for assessing frequency or probability because large classes are usually reached better and faster than instances of less frequent classes. However, availability is affected by factors other than frequency and probability.

These other dimensions of availability are sources of bias. They include retrievability, imaginability, and illusory correlation. The first of these concerns the proximity, in time or space, of various alternatives: those that are closest are available. Imaginability is self-explanatory. Finally, illusory correlation describes the human tendency to interpret evidence in the light of pre-existing expectations.

In the present context, the significance of availability is that it may inhibit analysts from forecasting outcomes that are dissimilar to those that they have experienced.

II.2.6 Representativeness

The third of Tversky and Kahneman's simplified decision rules is the representativeness heuristic. When this is used, the probability that case A came from group B is evaluated by the similarity between A and B. When judgements are made on this basis, prior probabilities tend to be ignored. For example,

consider the assessment of whether a country drawn at random is likely to try to restructure its external debts. The base-rate frequency of restructuring is relevant to this, yet an analysis based on the representativeness heuristic would ignore base-rate frequency and produce a biased estimate.

If man as an intuitive statistician disregards or misuses prior probabilities, he does not follow the correct Bayesian rule. Edwards (1968) maintains, on the basis of experimental evidence, that humans act, to an extent, in a Bayesian manner: that is, prior probabilities are reflected in assessments of evidence. However, Edwards goes on to conclude that humans fail to absorb the full impact of evidence: i.e., they are conservative. Further, Slovic and Lichtenstein (1971) argue that Edwards's evaluation of human performance is over-optimistic.

The representativeness heuristic has a number of additional dimensions. First, analyses based on it tend to overlook **sample size**. Sample variability falls as sample size increases, and a reliance on representativeness may result in biased estimates, generated by small samples. This effect has been called the 'law of small numbers'. A common manifestation of this is the human predisposition to see regular patterns in random processes. For

example, gamblers are prone to believe that fair roulette wheels or lottery-draws exhibit non-random and regular sequences. Tversky and Kahneman (1974) have termed this the "gamblers' fallacy". For example, it is evident in the beliefs of chartists concerning the behaviour of the prices of stock market securities, in the face of a considerable body of evidence that fails to reject the efficient markets hypothesis in its weak and semi-strong forms (Elton and Gruber, 1987: ch.15).

Secondly, there arises the question of **predictability**. Predictions should be made only if the information that is available warrants this action. If information is absent, or is being ignored, then predictability is nil.

During the early 1980s, bankers appear to have been slow to change their views about countries that on any objective assessment were becoming poor credit risks. For example, the Institutional Investor's rating of Mexico fell from 69.0 in September 1981 to 54.8 one year later. Nevertheless, that fall still left Mexico rated ahead of Algeria, Portugal and Thailand. The survey on which the September 1982 rating was based was carried out in mid-summer, shortly before the Mexican financial crisis broke in August. A possible explanation of this is that

bankers' views were formed, not on the basis of objective information pertaining to credit-worthiness, but on the basis of general sentiment and opinion about individual countries.

Thirdly, there is an aspect of representativeness that is known as the **illusion of validity.** This refers to the human tendency to have more confidence in predictions when the predictors are internally consistent. The problem is that internal consistency may merely reflect redundancy: that is the case when predictors convey essentially the same information and are highly correlated. Given a set of valid predictors, predictions based on them as a group will surpass in accuracy predictions yielded by a subgroup only when the variables are statistically independent. Tversky and Kahneman (1974: p.9) note that

> redundancy among inputs decreases accuracy even as it increases confidence.

The final dimension of the representativeness heuristic is the tendency of analysts to overlook **regression towards the mean.** This phenomenon conflicts with a belief that predicted output should be maximally representative of input. If an analyst holds to this latter belief, then he will tend to give an undue weighting to extreme values among the predictors.

II.2.7 Heuristics and biases: summary

In summary, each of Tversky and Kahneman's heuristics is a possible source of bias in human information processing. The heuristics do not obey normal statistical principles, and the use of any or all of them may cause systematically biased forecasts. Moreover, another characteristic of human behaviour may compound the problem. If the analyst suffers from hindsight bias, then he may consistently fail to revise his decision processes in the light of adverse outcomes. Fischoff (1980: p.341) considers that

> In hindsight, people consistently exaggerate what could have been anticipated in foresight. They not only tend to view what has happened as inevitable, but also to view it as having appeared 'relatively inevitable' before it happened...they even misremember their own predictions so as to exaggerate in hindsight what they knew in foresight.

II.3 Heuristics and biases: evidence

The evidence to support the existence of judgemental biases comes both from experimental work and from studies of actual decision processes, and is drawn from many areas of human decision. Taffler (1984b) warns of potential dangers in generalising from one task environment to another, or from laboratory research to real decisions. However, as he points out, research evidence suggests that biases are

> not task associated but are part of our personal information processing characteristics and that the same underlying behaviour patterns exist both outside and inside the laboratory (Taffler, 1984b: p.7).

Libby and Lewis (1982) survey the literature on human information processing in accounting up to that date, and conclude (p.273) that

> lens model research suggests that accountants and others may not be as proficient at certain aspects of decision making as was once thought. Inaccuracies arise from inconsistency in the application of decision rules, and misweighting of evidence - the latter resulting from the use of heuristics.

Taffler (1984b) focuses on biases in investment analysis and fund management, but he also quotes evidence to support the hypothesis that biases of judgement are a characteristic of human information

processing, and are not specific to any particular area of decision.

In any field of activity, formal quantitative forecasting models provide explicit estimates of forecasting error. Although the existence of forecasting error may also be recognised by analysts who use less formal approaches, it has been suggested by Slovic, Fischoff and Lichtenstein (1980) that humans suffer from a general tendency towards over-confidence in the precision of judgemental forecasts. Casey and Selling (1986) have presented evidence confirming this hypothesis, drawing on an experiment in bankruptcy-prediction.

A sceptical viewpoint on the 'bias' literature is contained in Berkeley and Humphreys (1982). First they argue that not all heuristics are necessarily suboptimal or inefficient. Moreover, they suggest that there is a certain circularity in the way in which the evidence is frequently examined: where a heuristic produces a correct decision, they suggest that protagonists of the 'bias' viewpoint will mistakenly ascribe it to chance, or to the use of a heuristic that gives good results in this 'special case' but not generally.

Secondly, Berkeley and Humphreys explore a 'bias

heuristic' whereby experimental design is 'anchored' on the existence of bias, and then employs 'available' tasks and subjects, and treats the results as 'representative' when generalising from it.

Berkeley and Humphreys looked at a large sample of published works that cited Tversky and Kahneman (1974), for the period 1975-1980. Inter alia, they found that these papers overutilised students in their experiments and that there was a low proportion of case studies. As against this, of 15 experiments concerning the use of heuristics in accounting that are reported in Libby and Lewis (1982, tables 1,2 and 3), nine utilised practitioners, and of the remaining six, three utilised graduate business students exclusively. Thus in these experiments at least, which on the whole confirm the existence of bias, the participants may be regarded as well informed and working in their own areas of specialisation.

II.4 Judgemental biases: implications

The acuteness of the problem of 'heuristic biases' falls as we move from fully qualitative approaches towards more formal analyses, although the potential for bias can seldom be eliminated completely.

Einhorn and Hogarth (1982) have observed that, while it is well established that correlation does not necessarily imply causation, it is not always recognised that causation does not necessarily imply high univariate linear correlation. They conclude from the latter point that sole reliance on statistical measures is insufficient for understanding data. Moreover, As Kahneman and Tversky (1979: p.414) put it

> Opinions and intuitions play an important part even where the forecasts are obtained by a mathematical model or a simulation. Intuitive judgements enter in the choice of variables that are considered in such models, the impact factors that are assigned to them, and the initial values that are assumed to hold.

This list may be extended in at least three ways. First, economic and financial data will be 'noisy', and also subject to revision. Moreover, there will frequently be a 'missing values' problem. These points apply equally to the training sample and to

the input data for out-of-sample forecasting, and there is clearly scope for human intervention to cope with them.

Secondly, human judgement will determine how frequently models are to be tested for stationarity and re-estimated. In these contexts, Kahneman and Tversky (1979) suggest that human intuitions and judgement form an irreducable component, and that the key question is how to de-bias and improve the valuable information that they convey. Formal statistical tests have an evident role here.

A third role for human judgement in forecasting concerns the interpretation of results. All forecasts of financial and economic events are probabilistic. Formal models make this explicit, with explicit presentation of forecasting errors and confidence intervals, but there is still room for human intervention. Model forecasts must be interpreted in the light of information that is not captured ex ante by the model - and that probably could not be so captured by any conceivable statistical model - concerning, for example, exogenous shocks to the system such as a trade embargo, a war, a drought, the formation of an effective producers' cartel.

II.5 The case for quantitative methods

In general terms, the fully quantitative approaches allow a number of sources of bias to be avoided. Because models need not employ an incrementalist approach to a problem, the 'anchoring and adjustment' heuristic may be avoided. A satisfactory model will form its predictions solely on the basis of the values of predictor variables, and thus the pitfalls of the 'availability' heuristic will be avoided. Finally, by various means, the 'representativeness' heuristic may be avoided also. First, the data may be 'Winsorized' to bring outliers in towards the centre of the distribution. Secondly, if a large sample is not available, predictions will be offered on the basis that they incorporate characteristics associated with small samples. Thirdly, multicollinearity among predictors can be tested for. Fourthly, feedback can take place, from systematic forecasting errors, to re-specification and re-estimation of the model.

As we have seen, in areas where human judgement and intuition are unavoidable, statistical techniques may be used to minimise bias: these areas include the specification of the model and the interpretation of results. This minimum bias may exceed zero, but even

if this is the case, a formal model will arguably out-perform a human who bases his predictions on the same data. The hypothesis here is that a formal model can eliminate the random error in human judgement, and thus apply human judgement consistently. According to Fischoff (1980: p.337)

> Two decades of policy capturing studies persistently [concluded that] (a) simple linear models, using a weighted sum of the cues, did an excellent job of predicting judges' decisions; (b) the judges claimed that they were using much more complicated strategies.

This technique of 'modelling man' is known as bootstrapping. If the underlying 'man' is subject to heuristic biases, then bootstrapping should capture them. In this case an appropriate research strategy would be to seek a better model. However, whether or not human judgement in a particular case is subject to heuristic biases, it is clear that in general a strong case exists for the use of quantitative models.

Procedures such as discriminant analysis and linear regression utilise optimising methods in the estimation of coefficients, to produce 'proper linear models' of the relationship between cues and criterion. Linear models whose coefficients are not based on optimising this relationship are known as 'improper': examples include 'equal weighting' models, models with randomly selected coefficients,

and paramorphic models whose coefficients are chosen, by optimising methods or otherwise, so as to attempt to simulate the actions of human decision takers.

Hoffman (1960) borrowed the term **paramorphic** from mineralogy, and used it to denote the representation of human judgement by a mathematical model. Bowman (1963), Goldberg (1970), and Dawes (1971) compare human judges with their paramorphic representations, and find that the latter consistently out-perform the former in the chosen task settings. Kaplan and Urwitz (1979) report similar results for the bootstrapping approach, when applied to judgementally-derived bond ratings using regression. Their findings are particularly relevant to chapter XII below, which explores ratings of country risk.

Goldberg suggests that the performance of paramorphic models follows from their elimination of the random variation in human judgement. Dawes and Corrigan (1974) investigate whether the performance of paramorphic models, in terms of predicting the criterion variable, is superior to that of other linear models. They report that equal weighting (in fact unit weighting) models perform about as well as regression models, in five experimental settings. Both perform better than paramorphic models, while random linear models perform about as well. Dawes

(1979) suggests that the 'model of the judge', and other improper linear models, are nearly optimal, because proper linear models typically have 'flat maxima' and so are robust to departures from optimal coefficient values. Moreover, equal weighting models have the advantage that they are free from sampling or measurement error in estimation. However, Einhorn and Hogarth (1975) conclude that if the sample size exceeds 200 and the criterion variable is well defined, a regression model (i.e. a 'proper linear model') is preferable.

This evidence suggests that a linear model will be at least as good as human judgement, provided that (i) the model is correctly specified and (ii) there are no environmental nonlinearities in the relationship between cues and criterion, that the judge takes account of. On the first of these points, Einhorn and Hogarth (1975) suggest that choice of variables is more important than the size of the weights placed on the cues.

The second of these points, concerning non-linearities, is discussed in Libby (1976a, 1976b), Goldberg (1976), and Schepanski (1983). Libby found that a linear discriminant model of human judgement, in an accounting setting, was outperformed by the judges that the model was attempting to

bootstrap, while Goldberg found that nonlinear transformations of the cues produced a model that outperformed the judges. Libby concludes from this that the judges were employing accurately a nonlinear rule, and that to replace nonlinear decision-makers by linear models of them could therefore be mistaken. Schepanski tested several representations of information processing behaviour in credit judgement. He finds that a linear model may not be appropriate - that in fact a nonlinear model provides a better representation of human judgement in the cases examined.

The second point (b) made by Fischoff, concerning judges' self-insight, has recently been tested experimentally by Mear and Firth (1987). They find that in making risk and return predictions for securities, expert judges' indications of cue importance were consistent with the variables found to be significant in regression models of their judgemental behaviour. This indication of self-insight raises a question mark against Fischoff's second point, but not against the case for quantitative models.

II.6 Country and corporate risk: quantitative approaches

II.6.1 Modelling approaches to country risk

Among the earliest multivariate investigations of country risk in the recent period is that of Frank and Cline (1971), in which linear and quadratic discriminant models were developed. Since then, the literature has expanded in parallel with the deepening debt crisis. Linear discriminant analysis and logit analysis have been the most widely use statistical techniques, while principal components analysis, cluster analysis and the analytic hierarchy process have also been used. Surveys of the literature up to the early 1980s may be found in Saini and Bates (1984), Somerville (1985), and Heffernan (1986). Recent contributions include Morgan (1986), Berg et al (1988), LLoyd-Ellis et al (1989), and Johnson et al (1990).

These models vary widely, in terms of the statistical methodology used, the specification of the predictor variables, the spatial and temporal coverage of the estimating samples, and the sources of data. Schmidt (1984) uses discriminant, logit and cluster analyses, but his paper is unusual in applying more than one

technique. This thesis will attempt to provide a general model, in several ways. First, it will use both of the most widely used statistical techniques (i.e. discriminant and logit analyses), augmented by cluster analysis and automatic interaction detection (AID). Secondly, the predictor variables will be drawn from a very large initial set, using interactive stepwise techniques. Thirdly, the data base will cover all of the significant debtor LDCs, for the period from 1979 to 1989. Fourthly, the data benefits from being processed by the Economist Intelligence Unit, where it has beeen modified, where appropriate, in the light of other information known to country experts.

II.6.2 Corporate risk

Logit analysis and discriminant analysis have been used to investigate corporate bankruptcy. However, the connection between research on corporate and sovereign risk is not limited to the statistical methodology. The analysis of country risk is concerned with the ability of countries to service their external debts: this is the macroeconomic counterpart to the microeconomic problem of corporate insolvency. In both cases, financial ratios are the building blocks of the multivariate models. In the

corporate context, the statistical techniques have traditionally been discriminant analysis and logit and probit analysis; recently, recursive partitioning has also been applied. The literature on predicting corporate distress is surveyed by Barnes (1987), while other recent contributions include Frydman et al (1985), Gombola et al (1987), Lau (1987), Moses and Liao (1987), Srinivasan and Kim (1987), Aziz at al (1988), Hopwood et al (1989), BarNiv and Hershbarger (1990), Gilbert et al (1990), and Gupta et al (1990). Doukas (1986) specifically compared banker judgement of corporate solvency with multivariate statistical models, using Canadian data. His findings are of relevance to this thesis, in that he found that the models were unable to out-perform the banks. However, these results are based on a sample of only 18 banks, which is rather small for general conclusions to be drawn.

II.6.3 Political risk

The risk that is borne by the external creditors of less developed countries inescapably includes an element of political risk. For this reason alone, it is clear that multivariate models that are based essentially on financial and economic data may require judgemental interpretation - unless, that is,

the elements of political risk are always successfully captured by the model. However, the latter possibility is unlikely: according to Ascher and Overholt (1983, p.121)

> The most obvious distinction between political behaviour and most other behaviours subject to forecasting is the greater relative importance of single, discrete events in setting the direction of political developments. Elections, coups d'etat, wars, and policy choices are important per se, rather than solely as incremental additions to a larger pattern.

According to de la Torre and Neckar (1988), there has been a tendency since the early 1970s for political risk assessment in international firms to become increasingly systematized and formalised. Assessment will typically take place internally, at varying levels of organisation and sophistication (Mascarenhas and Sand, 1985), and it may also use the services of country rating agencies such as BERI, Political Risk Services, The Economist Intelligence Unit, and Data Resources Inc. These will be discussed further in chapter XII. At this stage it is sufficient to note that these rating systems explicitly include factors such as the operation of the political system, the degree of enfranchisement, civil war risk, and racial and nationality tensions (ICRG, 1984; EIU, 1988b). First of all, the assessment and rating of many of these political factors involves a considerable degree of judgement.

Secondly, according to de la Torre and Neckar these 'expert' assessments are:

> typically obtained as the end product of a multi-stage consultation process that may or may not involve Delphi methods. Some of these reports might include econometric data as well, but their major characteristic is the progressive ranking of a large number of countries according to a more or less explicit logic of analysis.

Rice and Mahmoud (1990) have surveyed the methods of political risk assessment that are in use by 105 Canadian firms. They have found personal judgement to be the most widely used method of forecasting, as Table II.1 illustrates.

Table II.1: Political risk forecasting methods: survey by Rice and Mahmoud

Method	Percentage of 105 firms that used method	Mean success score +
Personal judgement	79	2.58
Structured qualitative	23	2.04
Standardised checklist	27	1.96
Investment models	14	1.87
Scenario development	30	1.83
Statistical analysis	28	1.83
Delphi technique	7	1.29

Source
Rice and Mahmoud, 1990

Note:
+ The success score was as follows:
1: used, with no success;
2: used, with moderate success;
3: used, with great success.

Only 18 per cent of the sample used the services of one or more of BERI, Business International, and Frost and Sullivan (now Political Risk Services). According to the survey, statistical analysis was the third most widely used method, but well behind judgement in popularity.

The respondents to the survey reported their own perceptions of the various forecasting techniques. According to a statistical test quoted by the authors, the mean score of judgement was significantly greater, and that of Delphi less, than those of the other six (including statistical analysis) - which were equally ranked. By regressing the success rate on dummy variables representing different types of firm, the authors found that the success of statistical analysis was significantly (and positively) related to its use in financial institutions, while the success of judgement was significantly (and negatively) related to size of firm. However, this evidence on the success rate should be viewed with caution. The phenomenon of 'illusory correlation' tends to lead human judges to interpret evidence in the light of pre-existing expectations, with the consequence that the self-reported success rate is likely to be biased upwards.

Ascher (1982: p.234) has cautioned against the use of quantitative measures of political risk as being 'the antithesis of complexity and configuration', and he gives five reasons for this. The first reason is aggregation: these ratings condense or conceal information. The second is spurious generality: the rating model is applied to different countries, without discrimination. The third is over-simplification. The fourth is the (alleged) necessity to use only quantifiable factors: Ascher states that

> often the 'intutitive' awareness that one factor is relevant to another factor in a specific case cannot be translated into an explicit formulation.

Finally, Ascher suggests that time is spent on quantitative analysis at the expense of 'refined analysis'.

The first and third of these points apply with as much (or as little) force to formal quantitative models in any context. The second and fifth are criticisms of practice rather than principle. This leaves the fourth point, and the only direct test of this is the track-record of the alternative approaches to forecasting.

However, an indirect test may be made through analysis of the ratings themselves. Fischoff (1980)

has been cited (section II.5) on the good empirical performance of bootstrapping linear models, versus the performance of underlying judges who 'claimed that they were using much more complicated strategies'. In the context of political risk, expert ratings have been investigated by Mumpower et al (1987). They arranged for 51 political analysts to judge the political riskiness of 49 countries. Using regression analysis, they found that 75 per cent of the variance in the analysts' ratings could be accounted for by a linear combination of three variables: exchange rate differential, estimated inflation rate, and infant mortality rate (which the authors use as a measure of the level of socio-economic development). Thus perceptions of political risk are, according to Mumpower et al, significantly related to economic factors. The question arises whether, in general, political risk is orthogonal to economic risk, and this will be explored further in chapter XII.

II.7 Man-model interaction: the evidence

Klein (1981) reports extensive comparisons between econometric forecasting models and judgemental forecasts for the macro-economy of the United States. He observes that, inter alia, the judgemental forecasts were as accurate as those of the models in the short run, but that the models outperformed judgemental forecasts in the long run.

If quantitative models are to be used, a key issue is the way in which forecasts are to be produced: that is, whether the raw model forecasts should be modelled by human intervention and judgement ('man-model forecasts') or whether the model should be applied mechanically. In the early days of econometric modelling, Klein (1947: p.111) wrote that

> Statistical models...are not proposed as magic formulas which divulge all the secrets of the complex real world...Obviously any qualitative information...must be taken account of in order to make a proper forecast.

In 1981, after years of involvement in large-scale econometric modelling, Klein reported evidence to support that view. After the comparisons cited above between judgemental and model forecasts, he goes on to report (Klein, 1981: p.43) that

> the different man-model forecast teams do not differ markedly among themselves, but they do have a record that is far superior to that compiled by mechanistically applied models.

Klein observes that the similarity between teams persists over time even as the personnel change, which suggests that, unlike pure judgemental forecasts, man-model forecasts are not ad hominem.

Makridakis (1986) argues that no study of forecasting methods has shown a clear superiority for one method over another. In particular, he argues that judgemental forecasts have not been shown to be more accurate than those of quantitative models. In a recent survey, Makridakis and Wheelwright (1989) argue that the advantages of quantitative forecasting methods over judgemental methods include consistency and cheapness. However, quantitative methods cannot predict changes in established patterns, nor distinguish whether deviations from patterns are temporary or permanent. Judgemental methods may do so, but the problem of judgemental bias arises at this point. Judgemental methods may also make more use of available information than quantitative techniques, but again the problem of judgemental bias arises. Since human judgement is involved in the specification of quantitative models, and in the choice and interpretation of statistical tests, bias is a factor in both types of model.

The conclusion reached by Makridakis and Wheelwright (1989, p.311) is that

> Quantitative methods and judgement are complementary and must be integrated to produce accurate and usable forecasts.

III Approaches to crisis resolution

III.1 Introduction

The purpose of this chapter is to develop a definition of the event to be predicted: that is, the dependent variable for the quantitative analysis. This raises difficult questions of definition and measurement. The conclusion will be reached that the primary focus should fall on the emergence of arrears of debt-service, but to reach this point it is necessary to explore the reactions of debtors and creditors to the onset of difficulties in meeting debt-service commitments.

Section III.2 raises general issues concerning the enforcement of international debt-servicing obligations. Section III.3 outlines the events that may follow the onset of debt-servicing difficulties. Sections III.4 to III.6 describe recent approaches to restructuring debt-service payments, while III.7 is a résumé of recent innovatory developments. Section III.8 looks at candidates for the role of dependent variable in the statistical models of later chapters, and concludes that the primary event of interest is the emergence of arrears. The final section, III.9, is concerned with sources of data.

III.2 Sovereign and corporate debtors compared

III.2.1 Legal issues

In the industrialised countries, the creditors of a defaulting corporate debtor have access to legal redress. For example, in the United Kingdom, failure by a company to comply with a court order to discharge its liabilities may result in its being placed in receivership by the court. In the last resort, a compulsory liquidation may be ordered.

International sovereign debts cannot normally be enforced by such methods. The courts of the debtor country may not be independent of the government. Political disturbances may make the courts inaccessible to foreign creditors, even if the courts themselves are independent. Even if a creditor succeeds in obtaining a judgement against the debtor, it is likely to be rendered in the local currency, and this may be inadequate to discharge the hard-currency debt.

Alternatively, an external creditor may seek redress externally, for example in the courts of his home country. According to Nichols (1984), in the period before the second world war, holders of defaulted

government debt were almost totally unable to avail of judicial remedies, in part because the doctrine of sovereign immunity prevailed generally in the creditor countries. Since then, worldwide legal changes have reduced the immunity of sovereigns from suit. For example, in the United States, the Foreign Sovereign Immunities Act of 1976 provides that subject to limited exemptions and existing international agreements, a 'foreign state' (defined to include the state, any subdivision of it, and a wide category of state agencies) is immune from the jurisdiction of state and federal courts. (United States Code, 1982 ed.: p. 279). However, a foreign state may waive its immunity, and according to Nichols (1984), bank lenders obtain waivers in most cases. In the United Kingdom, the State Immunity Act 1978 provides, according to Mann (1984: p.115):

> a residual immunity which a foreign State can claim in relatively few cases.

These develpoments are of limited significance. Even if judgement is obtainable by the creditor, the problem of enforcement will remain. Defaulting governments may have few external assets, and assets may be moved beyond the jurisdiction of the court, to forestall attachment. Moreover, it is unlikely that a sovereign state could be forced to liquidate or transfer ownership of real assets to pay creditors.

In the case of international private debts, the problem of sovereign immunity does not arise. However, the resolution of a default is nevertheless likely to be more complex than in the case of purely domestic debts, because all the other difficulties remain: namely, the degree of accessibility and independence of the courts of the debtor's country, and the convertibility of the currency.

III.2.2 The right of set-off

Where a defaulting debtor has bank deposits with a bank creditor, the bank may attempt to set off these deposits against the debt. The laws of the jurisdiction within which the bank is based will determine whether the bank has the right of set-off, and if so, how it may be exercised. Commonly, a condition of set-off is that the obligations at issue must be between the same parties acting in the same capacities. This is not usually true in the case of sovereign debts and credits, which typically involve different government agencies respectively. It would then be for the courts to decide whether the bank had the right of set-off. In the absence of outright repudiations in the recent past, the right of set-off has not been a pressing issue.

III.3 The emergence of debt-servicing problems

The first manifestation of servicing problems on an international debt is generally the emergence of arrears of debt service. A moratorium is a further stage, whereby the debtor formally states his inability, or refusal, to continue with regular service payments for the time being. Recent examples of moratoria are Peru (1985) and Brazil (1987). In the case of Peru, the government decided unilaterally to limit debt-service payments to ten per cent of export earnings. In the Brazilian case, the government suspended service payments on its bank debt. However, the process of resolving debt-servicing problems has in many cases begun without the debtor formally declaring a moratorium.

Outright repudiation takes matters one stage further, to a conclusion that has not been reached by any country during the years covered by this study. Since the purpose of debtor-creditor negotiations is to achieve a resumption of service payments, and also a resumption of access by the debtor to external capital, all parties normally have an interest in negotiating. Generally, it is in debtors' interests to retain access to the international capital markets and to avoid the exclusion that would follow

repudiation. In the interwar period, creditors of sovereign debtors were largely private bondholders. This is not now the case, and repudiation of bank debt would imperil short-term trade financing.

According to Cline (1983), the only recent cases of outright repudiation have been those by North Korea and Cuba. This being the case, creditors can and do have a degree of leverage over defaulting debtors. A country that is not prepared to repudiate its debts will ultimately have to negotiate a restructuring of its debt-service payments with its creditors, and normally the creditors - whether they be commercial banks, or the 'Paris Club' of official creditors - will require a country to adhere to an IMF adjustment programme, before they will agree to restructure.

In certain jurisdictions, the incentive to creditors to negotiate may be reinforced by the attitude of the supervisory authorities of the commercial banks to the emergence of arrears. For example, banks in the United States may not generally continue to treat interest as accruing when it is beyond 90 days in arrears.

However, these pressures towards negotiation are offset in a number of ways. First, on the creditor

side there is generally not a single entity - instead there is usually a sizeable group of banks or official creditors. For example, in July 1989, Mexico's bank creditors numbered about 500. The interests of these creditors may well be in competition, particularly in the case of banks. Secondly, free-rider problems tend to occur: restructurings typically include the provision of 'new money', which allows, inter alia, service payments to continue on existing debt. Clearly, any creditor bank would benefit from the implementation of a 'new money' agreement to which the bank itself was not a party. Generally, agreements do not take effect until some 'critical mass' of creditors has adopted them - and this may take a considerable period of time. Thirdly, the recent strengthening of loan-loss provisions has reduced the level of the threat posed by the debt problem to the lending banks. This certainly reduces one source of pressure for the banks to become involved in conventional restructuring arrangements; however, it may also facilitate longer-term solutions to the problem, involving debt forgiveness or debt consolidation on concessional terms.

Correspondingly, there may be obstacles to negotiation and agreement (albeit of a different nature) on the debtor side. No agreement will be

possible unless both sides benefit. Therefore any agreement may meet resistance within the debtor government or from the public or the political opposition - because there will be a tendency to discount the domestic benefits of any agreement, relative to the benefits perceived to accrue to creditors.

An equally potent obstacle to agreement, on the debtor side, arises from the usual requirement, imposed by creditors, that debtors implement specific changes in economic policy - usually by way of adherence to an IMF adjustment programme. However beneficial such policy changes may be in the medium term, there will always be sharp short-term adjustment costs. Moreover, adherence to an externally imposed programme will probably be seen as involving an erosion of political sovereignty, and be resisted for this reason also.

Griffith-Jones (1989), addressing the question why unilateral action by debtors has in fact been comparatively rare, concludes that the major reason has been the uncertainty of the impact that such action would have on the international banking system. She concludes that the prime factor is fear of the impact on world trade flows, rather than the risk to an individual debtor's own trade credits.

III.4 Approaches to debt restructuring

According to the International Monetary Fund (1985: p.31), debt restructurings consist of

> rescheduling or refinancing of debt-service payments in arrears and/or of future debt service payments, undertaken in response to external payments difficulties.

Rescheduling involves formal deferment of debt-service payments, with new maturities applying to the deferred amounts, while refinancing consists of either a rollover of maturing debt obligations or the conversion of existing or future debt service payments into a new medium-term loan.

Refinancing in this context is to be distinguished from voluntary refinancing, whereby countries that have free access to international capital markets will take advantage of favourable borrowing conditions to retire old debt, and replace it with new debt carrying more advantageous terms. For example, Thailand carried out such a voluntary refinancing in 1986 (EIU, 1986: p.27). In contrast, this thesis is concerned with involuntary refinancing, as defined within the **Debtor Reporting System** (DRS) of the World Bank, where a country does not have free access to the capital markets. Refinancing is an alternative to rescheduling,

implemented by agreement between creditor and debtor, in order to maintain debt service flows.

International debts may be owed to creditors in one or more of three categories: multilateral institutions (e.g. the World Bank), other official creditors, which are usually referred to simply as 'official creditors' (e.g. governments), and commercial lenders (e.g. banks). Debt owing to multilateral institutions has never been restructured, and this section is concerned only with creditors in the other two categories.

Typically, there is a group of creditors, rather than one only: for example, Mexico had about 500 bank creditors in summer 1989. The creditor group is normally smaller for official debt. However, in the case of bank debt there is normally a relatively small group of major creditors, with the majority of banks having a relatively small involvement.

Because of the multiplicity of creditors, negotiations on restructuring are normally conducted multilaterally. This is more efficient than negotiating with each creditor separately, it ensures uniformity of treatment for creditors, and it avoids free-rider problems (see III.3).

III.5 The restructuring of official debt

Usually, the negotiations leading to the restructuring of official debt take place within the framework of the Paris Club. This is an informal committee of official creditors that meets in Paris, at the request of the debtor government, under the chairmanship of an official of the French finance ministry.

In 1976, the sole official multilateral debt restructuring was the agreement reached between Zaire and its creditors. The number of of such agreements rose fairly steadily to a peak of 21 in 1985, of which all but four were reached within the framework of the Paris Club. Sixteen agreements were concluded in 1986, and seventeen in 1987. In the years up to 1987, the total debt restructured grew to reach a total for 1987 of US$24.7bn. In 1988, there was a fall to 15 in the number of reschedulings, and the total debt rescheduled fell to US$8.6bn. However, the upward trend was resumed in 1989: up to September, official debts of US$15.9bn were rescheduled in 18 agreements. (World Bank, 1989-1990; IMF, 1988c). The largest single agreement on record during this period occurred in 1985, when Chile and its creditors restructured debts totalling

US$10.9bn.

The debt covered by negotiations within the Paris Club includes loans made by governments and their agencies, and commercial credits guaranteed or insured by governments or government agencies in the creditor countries. In most cases up to about 1985, liabilities of both the private and public sectors of the debtor country were covered in Paris Club negotiations, but since then there has been a tendency for private-sector debt to be excluded, at the request of the debtor country.

Typically, and contrary to the practice in restructuring bank debt, interest and principal obligations have both been restructured. As regards maturity, the basic approach of the Paris Club has been to include only debt that had an original maturity exceeding one year, although this has been breached on occasion. Another principle that has also been sometimes breached is that previously rescheduled debt should not be rescheduled in later agreements.

The consolidation period of an official debt restructuring is the period in which debt-service payments to be consolidated or rescheduled have fallen or will fall due. (In the case of bank debt,

consolidation refers only to principal.) The consolidation period is generally twelve to eighteen months, but it may be longer, and may take the form of a multi-year restructuring agreement (MYRA). Detailed definitions of this and other terms are set out in IMF (1988a: appendix II).

Since debt service problems are symptomatic of balance-of-payments difficulties, and these may result at least in part from inappropriate domestic economic policies in the debtor country, Paris Club members have sought assurance that countries seeking debt relief would make appropriate policy adjustments. This has taken the form of requiring the debtor country to have reached agreement with the IMF over the use of the IMF's resources, where the debtor is a member of the IMF. In the case of non-members, corrective policies have been agreed directly.

III.6 The restructuring of bank debt

When major borrowing countries encountered difficulties in servicing their bank debts in the early 1980s, no such forum as the Paris Club existed for conducting negotiations between debtors and bank creditors. The system that has emerged involves the establishment, in each case, of a coordinating or steering committee of a small number of banks. Membership is determined on the basis of exposure, and to obtain a regional balance, and the committee acts on behalf of and in liaison with all bank creditors.

In 1978, the bank debts of Peru and Jamaica were restructured. By later standards, the amounts were small: for Peru, a total of US$386m. in two separate restructurings, and for Jamaica US$63m. The number of restructurings rose thereafter, slowly until 1982, and then in 1983 it jumped to 18, with others under negotiation. The total amounts of bank debt restructured under agreements reached in principle were US$34.6bn in 1983 and US$105.1bn in 1984, and during these years 80.7 per cent of the total was accounted for by Argentina, Brazil, Mexico and Venezuela.

In the period up to December 1988, 1984 was the peak both for the total amount of bank debt restructured and for the number of agreements in principle (22).

In 1985, 14 agreements in principle covered amounts that separately and in toto were relatively small by recent standards. In each of the three subsequent years, the total exceeded US$60bn, and over two-thirds of the three-year total was accounted for by just four agreements, one for each of the four countries listed above, with each exceeding US$20bn.

Comprehensive accounts of recent restructurings of bank debt may be found in IMF (1986a, 1988a, and 1989a), and World Bank (1989-1990).

The relatively small scale of restructurings before 1981 reflected in part the limited extent of difficulties in debt servicing in that period. However, it also reflected the prevalent practice whereby banks rolled over outstanding debt obligations, or consolidated them into new medium term loans, on a bilateral basis and without engaging in formal restructuring. In contrast, since 1981 creditors and debtors have become more willing to engage in formal and multilateral restructurings.

In almost all cases, the category of debt covered by these agreements has included medium- or long-term bank debt incurred or guaranteed by the government of the debtor country or its agencies. Unguaranteed private sector debt was only covered in a limited number of cases up to 1983. However, in the following two years this category of debt featured in nearly half of all restructurings. Debt that was guaranteed by the government of the creditor country, or an agency of it, has always been excluded on the grounds that this is a species of official debt. The banks have tended to resist the restructuring of short term debt, not always successfully. For example, about half of all bank-debt restructurings up to 1983 covered short term debt.

Largely because of the regulatory procedures in their home countries, and also for reasons of accounting practice, banks have been unwilling to reschedule interest in arrears or future interest due. However, in recent years a number of countries have allowed arrears of interest to accumulate as a means of obtaining new financing, and in some of these cases the banks have in effect rescheduled interest payments.

Up to 1983, the consolidation period was typically between one and two years. In 1984 the first bank

multi-year rescheduling agreements (MYRAs) were negotiated, with Ecuador, Mexico, and Venezuela, and others have been negotiated since then. MYRAs involve a consolidation period that runs for more than two years beyond the date of signing, and are aimed primarily at eliminating a hump in amortisation payments.

Generally, the implementation of an agreement to restructure bank debt has been made conditional on reformed macroeconomic performance by the debtor country. In some cases, the debtor has been required to have an IMF programme in place, either before agreement could be reached in principle or before the agreement could be signed. Many agreements have been conditional on compliance with performance tests arranged by the Fund.

III.7 Handling debt servicing difficulties: recent innovations

III.7.1 Introduction

Since the 1970s, restructurings of bank debt have involved reschedulings and refinancings. This remains the case, but in recent years new instruments and techniques have also been utilised. These have included currency redenomination, secondary-market transactions and debt buy-backs, swaps and conversions of various types, securitisation, and relending and onlending. Relending and onlending are marginally relevant to this thesis, and so are not discussed further. They are described in IMF (1988a: p.56). The other new techniques and instruments are described in the following sections.

III.7.2 Currency redenomination

Currency redonomination allows existing debts to be converted into other currencies - for example, the domestic currency of a creditor bank, or the ECU, or a third currency. In some agreements, redenomination is at the discretion of the creditors, and in others at that of the debtor. The purposes of these

arrangements include currency diversification, the reduction of exchange rate risk, and the matching of debt service commitments more closely with foreign exchange earnings.

III.7.3 Secondary-market transactions

Since 1982, a secondary market has arisen in which bank debt may be traded. While the secondary market does not feature directly in restructurings, it is highly relevant to the restructuring process in that it provides a means of exit from the market for those banks that no longer wish to be involved. In recent years, exit has become a more attractive option, first because many banks have made substantial balance-sheet provisions against their loans to developing countries, and secondly because balance sheets among the creditor banks have, in general, been strengthening. When banks were effectively locked-in to lending to LDCs, it was easier to achieve their participation in restructuring packages than it is now, when an exit option is available.

Discounts in this market averaged 30 per cent for the fifteen heavily indebted countries at the end of 1986, 55 per cent in December 1987, and 60 per cent in December 1988. (IMF, 1989a: p. 36.) The total

volume traded has been growing: the IMF reports estimated gross flows of US$40bn - $50bn for 1988, compared with US$5bn in 1985-86. For comparison, at the end of 1988, the total external debt of the developing countries was US$1240bn.

Most of the turnover in the secondary market has involved the debts of five Latin American countries: Argentina, Brazil, Chile, Mexico and Venezuela. According to the World Bank (1988-89: vol. 1, p.xxiii) these countries accounted for about 90 per cent of turnover in 1987, and continued to dominate in 1988. The World Bank also reports that interbank transactions, undertaken for purposes of portfolio adjustment, form a major share of the total.

Debt buy-backs enable countries to repurchase and retire their debt at a discount. In 1987, in an amendment to the 1981 refinancing agreement with Bolivia, creditor banks agreed to a buy-back. This took place directly between Bolivia and the banks, with the Bolivian authorities naming a price (11 cents on the dollar), and the banks choosing the amount of debt that they would be prepared to sell at that price.

Since then, Mexico and Chile have also negotiated buy-backs, and others are under negotiation.

III.7.4 Conversions and swaps

Conversions and certain types of swap have a similar significance to straightforward sales of loans in the secondary market - to the extent, that is, that they are undertaken bilaterally by a bank with some other party. However, in recent years some of these techniques have also been used as elements in restructuring packages.

The term 'swap' has several meanings. Currency and interest rate swaps may be transactions between debtors, and as such are neither a means of exit for creditors nor a component of restructuring packages. They do not change the level of debt, nor the aggregate levels of service commitments. Their effect is to alter the risk characteristics of the liabilities of the individual debtors.

Asset-based swaps may involve exchanges of claims between lenders. These merely change the identity of creditors. Unless one of these is a resident of one of the debtor countries, they do not affect the overall levels of external indebtedness. Their purpose is to rearrange the composition of the creditors' portfolios. In contrast, debt-equity swaps involve the replacement in banks' asset portfolios of debt by equity in corporations in the

debtor country. This involves a reduction in debt, although not necessarily in the level of cross-border liabilities. This type of swap is also known as a debt-equity conversion.

A second type of conversion involves the purchase of bank debt at a discount on the secondary market, by non-banks. The seller of the debt (who may be the original creditor bank) receives the discounted secondary market price in hard currency. The buyer then redeems the debt at near par at the central bank of the debtor country, in local currency which may then be used to make direct investments or to purchase financial assets. The buyer may be foreign, or a resident who possesses external assets.

A third form of conversion consists of the exchange of debt for bonds. In February 1988, Mexico accepted bids from 95 banks, whereby the banks exchanged US$3.7bn in debt for US$2.6bn in new bonds.

III.7.5 Securitisation

Securitisation can mean the issue of securities by borrowers as a substitute for direct borrowing. Securitisation in this sense has featured in recent restructurings, which have included, for example 'new

money bonds' and 'exit bonds'. The latter involve a conversion of debt to bonds, with the proviso that the creditor bank be exempted from future requests for new money and restructuring. Since 'new money' is frequently used to finance the retiral of bank debt, the issue of bonds of either type may imply a conversion of bank debt into bonds.

Securitisation has other meanings which are not directly relevant in this context, including (i) the transformation of less marketable claims into a more marketable form and (ii) the emergence of innovative financial instruments.

III.7.6 The Brady initiative

In March 1989, the US Treasury Secretary, Nicholas Brady, announced that officially sponsored, but market-based, reductions of debt service payments should be an integral part of creditor policy (Brady, 1989 pp. 73 and 76). Since then, the IMF and the World Bank have pledged a total of $20bn in support of new lending to support debt reduction programmes under the Brady initiative, and Japan has pledged a further $10bn. By October 1990, Costa Rica, Mexico and the Philippines had benefited under the initiative, to the extent of $9.5bn in debt

reductions, and a further agreement with Venezuela was pending. All these reductions involve a market-based menu of options, and typically utilise exit bonds. (World Bank, 1990-1991: vol. 1, pp. 6-8).

III.8 The event to be predicted

This thesis is concerned with the prediction of difficulties in servicing external debt. In principle, a country may be classified, in any period, into one of two groups: countries that are adhering to their debt-servicing commitments, and those that are not. However, difficulties arise in distinguishing these groups in practice.

The key event is the cessation of debt service payments. This, rather than the event of restructuring, is what is to be predicted. Restructuring is undertaken in reaction to difficulties in meeting debt-service commitments, or in anticipation of them. Most studies in this area focus on the event of restructuring, which in effect is taken as a proxy for the event of the emergence of arrears. However, according to Saini and Bates (1984: pp.349-350),

> The choice of this dependent variable seems to have been inappropriate, because countries have options other than a formal rescheduling when they are facing debt servicing problems.

In the work surveyed by Saini and Bates, only Saini and Bates (1978) and Feder (1980) modified the dependent variable in any way, and only to the extent of distinguishing between voluntary and involuntary

reschedulings. Recent work has continued to focus on rescheduling: for example, Lloyd-Ellis et al (1989).

Errors in dating the onset of debt-servicing problems are liable to flow from the use of 'restructuring' as the triggering event. There is frequently a lag, of variable length, between the onset of problems and the emergence of an agreement between debtor and creditors. In extreme cases, the lag is so long that some cases are well within the sample period at the onset of problems, but are outside it at the time of restructuring. Alternatively, the converse may hold. Even with shorter lags, the problem may emerge close to the ends of the sample period.

In summary, misclassification errors will arise from the use of the restructuring date as a proxy for the onset of debt-servicing problems, and the approach taken here is to focus on the emergence of arrears. The dependent variable is categorical, the categories being:

(a) country-year-cases where arrears have emerged in the given year;

(b) country-year-cases where debt-service commitments were adhered to in the given year.

The objective is to forecast group membership at date t+1, given the values of a set of predictors at date t. In the text, the following terminology will be

used: if a country falls into group (a) in year t+1, then t+1 will be referred to as an 'arrears year' for that country. We now define a categorical variable ARS, which in these circumstances will take the value zero for the country for date t+1.

Because the statistical procedures require distinct groups, cases close to an 'arrears' case are deleted from the analysis on the expectation that they will share characteristics with 'arrears' and 'non-arrears' cases. On these grounds, country-year cases that fall one, two or three years on either side of an 'arrears' year are defineded to be 'weak-year cases', and ARS is assigned the value 1, 2 or 3 as appropriate. In a few instances (see III.9), cases beyond this horizon are also classified as 'weak', and ARS is then coded 4. All other country-year-cases that are known not to have arrears on debt-servicing are classified as 'non-arrears', and ARS is then coded 9.

When referring to an arrears case, that status is determined by the value of ARS at t+1, but the predictors for the case are dated t. The same holds for weak-year and non-arrears cases.

III.9 Sources of data

The information on arrears and restructurings is drawn from sources that include the World Bank (1988-1989; 1989-1990), the IMF (1983; 1985; 1986a; 1988a; 1988c; 1989a), and the OECD (1983; 1987), and the individual quarterly and annual country reports of the Economist Intelligence Unit (EIU). These sources were supplemented by press reports - drawn largely from The Economist and the Financial Times. Where a restructuring is reported without information on a date at which arrears emerged, then the date of agreement in principle (for bank debt) or of the agreed minute (for official debt) is used to determine the date of the 'arrears' year. Where no other information is available, the date of signing is used, although this can be much later.

This procedure was also used in the small number of cases in which countries have encountered difficulties, but have nevertheless serviced their debts in full up to the point of implementation of a restructuring agreement. An example of this is the restructuring agreement between the Chilean government and its bank creditors that was reached in February 1987 (EIU, 1987). Chile was not in arrears immediately prior to this date. Nevertheless, the

immediate effect on the participating banks of the implementation of the agreement was that their cash flows from Chile were reduced below their scheduled levels: spreads were reduced, and interest charges were moved from semi-annual to annual payment. Thus the effect on cash flow to the banks resembled what would have happened if Chile had unilaterally allowed arrears to emerge, albeit with different timing. For this reason, events such as the Chilean restructuring of 1987 are included within the scope of the definition of the 'arrears' variable ARS, even though the terminology is a little inexact.

Where an agreement was not implemented, then the information relating to it is nevertheless used as described. If, as has usually happened, a stillborn agreement was superseded by another agreement at a later date, then all the intervening years are treated as 'weak-years', even if this involves treating more years than usual than weak. For example, Honduras went into arrears in 1981, and after several false starts, finally reached a restructuring agreement with its creditors in 1987. In this case, 1981 is taken as the 'arrears' year, while the succeeding six years are all treated as weak.

The virtue of using the date of arrears rather than

that of restructuring is illustrated by the case of Argentina. Argentina's bank debt-service went into arrears in 1982. An agreement in principle to restructure was reached in January 1983, but a renegotiation was requested by the newly-elected government. A further agreement in principle was reached in December 1984, and the final agreement was signed in 1985. Using the procedure as described, the key date is 1982. If alternatively the chosen date was to be that of restructuring (in principle or in fact), then the agreements of 1983, 1984 and 1985 would each be a candidate. It is the contention of this thesis that none would be appropriate.

IV Variables and data sources

IV.1 Introduction

This chapter introduces the data that will be used subsequently. Section IV.2 sets out criteria for the inclusion of variables in the data set. Section IV.3 describes the contents of the primary set of ratio-scaled variables, and the sources from which they are drawn, while section IV.4 describes a set of categorical variables. In section IV.5 the ratio-scaled variables are submitted to various tests for univariate normality. Although in many cases the results of the several tests are discordant, it is clear that many of the variables are not drawn from univariate normal populations. This provides the motivation for section IV.6, in which transformations to univariate normality are sought, and in which the Winsorization of sample outliers is described.

IV.2 Choice of variables: earlier studies and a priori considerations

Somerville (1985), Saini and Bates (1984) and Heffernan (1986) survey recent research on country risk. The most commonly-used quantitative techniques have been discriminant analysis and logit analysis, and studies using them have employed a wide range of variables. For example, in eight papers cited by Somerville (1985: Table 4), 25 different variables were utilised. Some classification is clearly required here: it is necessary to look for the dimensions of economic structure, policy and performance that earlier researchers sought to represent by their choice of variables.

An eight-fold classification encompasses all the variables used by earlier researchers, and also accords with a priori expectations.

(i) The domestic financial environment, including indicators of fiscal and monetary policy and of inflation.

The budget deficit is a key variable, indicating the extent to which fiscal policy has an expansionary impact on demand. The greater the growth rate of demand, the more likely is a deterioration of the current account. Fiscal expansion must be financed,

and a growing budget deficit is likely to feed through into domestic monetary expansion, or foreign borrowing, or both - particularly when domestic capital markets are underdeveloped. Monetary expansion will expand demand, and will also have inflationary consequences. If the exchange rate is allowed to float freely, domestic inflation will cause it to depreciate. However, if this is resisted the currency will become overvalued, and the traded-goods sector will suffer. The consequences will include deterioration in the capital and current accounts of the balance of payments, and capital flight is likely to occur. Moreover, a high and uncertain rate of inflation has significant real costs: it reduces the efficiency of the price mechanism, increases uncertainty, shortens time horizons, and causes widespread distortions in the allocation of resources.

(ii) Indicators of the structure and burden of debt and debt service.

The magnitude of debt and debt service may be expected, a priori, to be of relevance to debt servicing capacity, provided that we measure them relative to some measure that corrects for the size of the economy (e.g. GDP) or ability to pay (e.g. exports). The structure of debt may be of significance - for example, its maturity structure, the proportion that is owed to official creditors

(and is probably therefore on easier terms than the commercial portion), the proportion that is owed by public or publicly guaranteed agencies. The traditional debt-service ratio (i.e. service payments on medium and long term debt/exports of goods and services) is an ex post construct and may not be suitable as an 'early-warning' indicator, even when modified to include in the numerator all debt service (including private and including short-term) plus repayments on short-term debt. An ex ante measure is needed, such as a ratio to be described below, which takes as numerator a forecast of debt service falling due.

(iii) Balance of payments indicators, including international reserves.

The current account position is relevant to ability to meet debt-servicing commitments, and can be monitored by various measures. These include the current balance itself (relative to a measure of size, such as GDP), reserves, import cover.

(iv) Strength and diversification of trading sector, and the openess of the economy.

This category is related to (iii): improvements in the underlying strength of the trading sector will feed into the current balance. Diversification (in terms of overseas markets and export products) provides protection against demand-side shocks when these relate to particular markets or products. The

level of openess has two dimensions. One arises naturally out of the structure of the economy: the endowment of resources, levels of comparative advantage, and real incomes. Insofar as measures of openess such as imports/GDP reflect these characteristics, they are not directly indicative of imminent difficulties in debt servicing. However, the second dimension of openess relates to the level of trade protection. Insofar as this tends to raise domestic costs of production and so to misallocate resources, it is highly relevant to the central issue. However, the data set does not contain measures of the level of trade protection, so measures of openess are proxies.

(v) Indicators of economic structure.

Some indicators in this category reflect 'supply side' characteristics of the economy, which influence its capacity to service its debts. Examples are the following ratios: domestic savings/fixed investment, fixed investment/GDP. This category includes indicators of sectoral shares, which are included for exploratory reasons.

(vi) Indicators of the size of the economy.

These indicators are included to test the hypothesis that a country's bargaining position, relative to its creditors, reflects its size. The indicators include population, labour force, real GDP, real total

external debt.

(vii) Indicators of economic growth.

The growth rates of GDP and industrial output are included as indicators of economic performance.

(viii) Indicators relating to the external environment, including the behaviour of creditors

This category includes the growth rates of world output and of real GNP in the industrial countries; the fiscal impetus imparted by the governments of the Group of seven countries; unit values of exports and imports for the LDCs; LIBOR.

Suttle (1989: p.22) has suggested that the factors that influence a country's ability to meet its foreign exchange commitments are:

> domestic financial policies, debt management policies, structural (or supply-side) policies, the external economic environment and the behaviour of creditors.

The categories listed above cover Suttle's listing - and go beyond it in that, domestically, they include environmental as well as policy variables. No doubt there are omissions - generally because of shortage of data. Notably, the data-set to be described below includes no exchange rate variable. This is because, for many country-year cases, the official exchange rate is irrelevant for many transactions, while data on secondary exchange rates are very poor.

IV.3 The data

IV.3.1 Sources

The source of most of the data for individual countries is the the Country Credit Risk Service of The Economist Intelligence Unit Ltd. The primary sources for this data base are the IMF, the UN, the World Bank, the OECD, the Organisation of American States, the Asian Development Bank, the BIS, the Oil and Gas Journal, and the Petroleum Economist. A detailed listing of sources is given in EIU (1988b).

The original EIU data set consists of 129 financial and economic variables, and contains records for 66 countries for the years 1978 to 1989.

IV.3.2 Choice of countries

Table IV.1 lists the countries for which the EIU collects data. Two-letter reference codings are shown: these are used in other tables in this thesis.

Table IV.1: Member countries of the EIU data set

AL Algeria	ID Indonesia	PH Philippines
AR Argentina	IR Iran	a Portugal
BD Bangladesh	b Iraq	a Saudi Arabia
BL Bolivia	a Israel	SE Senegal
BR Brazil	IV Ivory Coast	a Singapore
CM Cameroon	JM Jamaica	a South Africa
CH Chile	JD Jordan	KS South Korea
CN China	KN Kenya	SR Sri Lanka
CG Congo	a Kuwait	SU Sudan
CB Colombia	LB Liberia	SY Syria
CR Costa Rica	b Libya	TW Taiwan
DR Dominic. Rep.	MW Malawi	TH Thailand
EC Ecuador	MA Malaysia	TT Trin. Tobago
EG Egypt	MX Mexico	TN Tunisia
ES El Salvador	MC Morocco	TK Turkey
GA Gabon	NC Nicaragua	a UAE
GH Ghana	NG Nigeria	UR Uruguay
a Greece	PK Pakistan	VZ Venezuela
GU Guatemala	PN Panama	YU Yugoslavia
HD Honduras	PP Papua NG	ZA Zaire
a Hong Kong	PG Paraguay	ZB Zambia
IN India	PE Peru	ZI Zimbabwe

Notes:
a not used - see text
b not used - too many missing data

Nine countries marked [a] were excluded on *a priori* grounds. Three of these are high-income oil producers on the Persian Gulf (i.e. Saudi Arabia, Kuwait and the UAE). Two (Hong Kong and Singapore) are high-income industrialised east Asian city-states. A further two (Greece and Portugal) are members of the European Community and of the OECD. The eighth, Israel, is a high-income developed economy, with a special relationship with the United States. The ninth case, South Africa, has a unique

political and economic structure.

The purpose of this thesis is to develop models relating to less developed countries. The features listed for these nine countries so distinguish them from the generality of less developed countries that it was thought appropriate to exclude them from the sample.

IV.3.3 The problem of missing observations

The remaining data, for 57 countries, were examined prior to further analysis. Most of the series contained a 'missing value' coding for some country-year cases. Also, in certain cases the value shown was obviously in error, and had then to be treated as a 'missing value' coding. The general approach to missing values was to substitute an estimated value, computed in one of four ways, as appropriate.

First, where the available data showed a clear time-trend, interpolation or projection was used. For example, this method was used where data was missing for Population. Secondly, for variables that appeared to be fairly stable over time, or that were fluctuating about a constant value, arithmetic

averaging was used. This method was typically used to deal with missing values for such structural variables as sectoral shares in GDP and the shares of one or two major markets in total exports. Thirdly, it was sometimes the case that alongside a case of missing data, other variables were to be found from which the value of the missing data could be inferred. For example, there were cases of missing values for the growth rate of GDP, while the values of GDP were available for each year. Finally, the primary data sources were consulted, and data found therein was used where possible.

After the data-file had been modified in this way, it was found that in the case of forty-three variables for Libya and nineteen for Iraq, there were still missing values for all years. The missing Libyan series included such essentials as Public sector deficit/GDP, Total short term debt, Short term interest, Total interest, Effective interest rate and Effective maturity. Moreover, several important series had replacements for missing values interpolated on the basis of very limited information: thus for example, only four original observations were present for the period 1979 to 1986 for Annual average growth rate of consumer prices. Of the nineteen missing series for Iraq, none was absolutely essential to the further analysis.

However, a number of crucial series, including all those relating to debt service payments, were only completed by estimating from very limited initial information. For example, in the case of Public sector deficit/GDP five values were initially missing, and in the case of all debt-service and interest-payment variables only four years' data were originally present.

Given the poor quality of the data for these two countries, they were eliminated from the sample. The sample thus consists of the 55 countries shown in Table IV.1, other than those marked [a] or [b]. None of these countries suffers from an excessively large number of missing values, those remaining being those that could not be eliminated by any of the methods discussed above. The reason for this is that, typically, for a particular country and variable, data was absent for all years, or was present for too few years for interpolation, projection or averaging to be acceptable.

Of the twelve years' data in the sample, 1987, 1988 and 1989 were set aside for out-of-sample testing. Also, because the transformations to be constructed included year-on-year growth rates, the year 1978 (the first year of the sample) was effectively lost. After all this, the estimating sample consists of the

years 1979 to 1986 for 55 countries, or 440 country-year-cases in all. Of these, 102 are arrears cases, 200 are non-arrears, and the remaining 138 are weak-year cases. The status of countries is determined by the values of the categorical variable ARS for t+1 = 1980-1987.

IV.3.4 Choice of variables from EIU data-set

At this stage it was found that 29 variables exhibited a large number of missing values, relative to the reduced sample size of 440 country-year cases. Of the reduced total of 55 countries, 29 had no data in any year for Arrears of debt-service. Of the variables that suffered from this problem, this was the only one that left a serious gap in the data: in all other cases where the variable might have been required, a complete substitute series was available. Examples of this were: (i) the end-year inflation rate: five countries had no data on this, but the data was complete for the period average; (ii) growth rates of imports and exports: nine countries had no data for either series, but in each case a substitute was available.

The number of complete series, after treating missing values, where possible, as described above, was thus

100. For various reasons, not all were usable:

(i) First, 36 variables were the 'raw material' for other ratio variables in the data set: examples are Trade balance (which is also present as a ratio with GDP), and Official debt (which is also present as a ratio with total debt).

(i) Two variables relating to the petroleum market were deleted, and replaced by a categorical variable (see below).

(iii) Finally, 26 variables were not used directly, but were used to construct new ratios or were otherwise transformed (see below).

Thus of 100 available variables, 36 were chosen as they stood. All of these have been used in other studies of country risk, or may be justified on a priori grounds for further exploration. These variables are set out in Table IV.2.

The World Bank, the IMF, the BIS and the OECD have agreed to a 'core definition' of debt:

> Gross external debt is the amount, at any given time, of disbursed and outstanding contractual liabilities of residents of a country to non-residents to repay principal, with or without interest, or to repay interest, with or without principal. (International Working Group, 1988: p.19.)

Table IV.2: Initial variable selection

AGRP	Agricultural output/GDP	%
CARA	Current account balance/GDP	%
DCPI	Consumer price inflation (annual average rate)	%
DCPR	Growth rate of real private consumption	%
DDCR	Growth rate of domestic credit	%
DFIN	Growth rate of real gross fixed investment	%
DGDP	Growth rate of real GDP	%
DIND	Growth rate of real industrial production	%
DMN2	Growth rate of money supply	%
DPOP	Growth rate of population	%
EFIR	Effective interest rate on total external debt	%
EFMT	Effective maturity of total external debt	years
EXIM	Exports of goods and services/imports	%
INDP	Industrial output/GDP	%
INPS	Interest on total external debt/total external debt service	%
INPX	Interest on total external debt/exports	%
INPY	Interest on total external debt/GDP	%
INVR	Real fixed investment/GDP	%
LABF	Labour force	mn
MCOV	Import cover (i.e. Total foreign reserves/imports)	months
MGRO	Growth rate of real imports	%
MYRA	Imports/GDP	%
NBTT	Terms of trade	1980=100
OCTD	Official creditors/total external debt	%
POPN	Population	mn
PSBR	Public sector deficit/GDP	%
SERP	Output of services sector/GDP	%
SIRA	Savings/investment	%
TDPX	Total external debt/exports	%
TDPY	Total external debt/GDP	%
TDRA	External trade balance/GDP	%
TSPX	Total external debt service/exports	%
TSPY	Total external debt service/GDP	%
XDPD	Index of price competitiveness	1980=100
XGRO	Growth rate of real exports	%
XYRA	Exports/GDP	%

In Table IV.2, and hereinafter, 'total external debt' refers to gross debt in this sense: that is, it consists of gross long-, medium- and short-term debt owed by public, publicly guaranteed and private debtors to non-resident creditors of all categories. Total external debt service consists of interest and principal repayments on long- and medium-term debt, and interest on short-term debt.

IV.3.5 New variables constructed

A further 34 variables were formed as transformations of other variables, most of which were drawn from the EIU data set. These 34 were then added to the initial selection of 36. A few of the transforms involve variables drawn from other sources. One of these, the US wholesale price index, was drawn from the IMF's International Financial Statistics, and was used to deflate certain US$ nominal variables. Data for the numerator of the ratio DCPY was also obtained from this source. Finally, to construct PDSPX, one-year-ahead forecasts of debt-service payments were drawn from the World Bank's World Debt Tables.

The 34 transforms were selected on the basis of existing research in this area, a priori grounds, and to maximise the use made of the EIU data set. They

included year-on-year growth rates (annual percentage rate of change). The alternative possibility of using a multi-year average was considered, and rejected because of its impact on sample size. These transformations are set out in Table IV.3.

At this point we have the primary data set. It consists of 70 financial and economic variables, and contains missing values, of two types. First, in three instances a ratio could not be calculated because the denominator had the value zero: this occurred twice for PSBRG and once for DCPIG. Secondly, missing values for the numerator of DCPY in all years prevented calculation of this ratio for Taiwan. These missing values turned out to be of no importance in the analyses of later chapters: the variables DCPIG, DCPY and PSBRG were always very far from being entered by the various stepwise statistical procedures. After this had been established in preliminary runs, these variables were removed from the data set, in order to restore it to its full size of 440 country-year cases.

Table IV.3: Transformed variables constructed

BHPCA	International payments: - (balancing item on capital account/balance on current account)	%
CARAG	Growth of (current account balance/GDP)	%
COPC	New loan commitments/population	US$
DCPIG	Acceleration of consumer price index	%
DCPY	Domestic credit/GDP	%
DMN22	Growth of money supply, lagged 2 years	%
FRQPY	(Current account balance minus principal repayments on total external debt)/GDP	%
GDPX	Public external medium and long-term debt/exports	%
ILMAG	Growth of international reserves	%
LABFG	Growth of labour force	%
MCOVG	Growth of import cover	%
MLDPD	Debt outstanding to multilateral institutions/total medium and long term external debt	%
NARY	(International reserves minus total external debt)/GDP	%
NBTTG	Growth of terms of trade	%
PDSPX	One-year ahead projection of debt service/current exports	%
PSBRG	Growth of (public sector deficit/GDP)	%
RDPC	Real total external debt/population	US$
RTD	Real total external debt	US$mn
RTDG	Growth of RTD	%
RY	Real GDP	US$mn
RYPC	Real GDP/population	US$
RYPCG	Growth of RYPC	%
STPD	Total external short-term debt/total external debt	%
STPDG	Growth of STPD	%
STPY	Total external short-term debt/GDP	%
TDPXG	Growth of (total external debt/exports)	%
TDPYG	Growth of (total external debt/GDP)	%
TDSR	(Total external debt service plus total external short-term debt)/exports	%
TSPD	Total external debt service/total external debt	%
TSPDG	Growth of TSPD	%
TSPXG	Growth of (total external debt service/exports)	%
TSPYG	Growth of (total external debt service/GDP)	%
XPM12	Exports to two largest markets/exports	%
XPP12	Exports of first plus second principal export products/exports	%

IV.3.6 Environmental variables

Information on external, 'environmental', variables was obtained from IMF Sources, and these variables are listed in Table IV.4. In the cases of PM and PX, the values associated with a particular country-year case were drawn from the series for fuel exporters or non-fuel-exporters, as appropriate.

Table IV.4: Environmental variables

WOP	Annual growth rate of world output	%
ICY	Annual growth rate of real GNP: industrial countries	%
LIB3	London interbank offered rate on US$ three-month deposits	%
RLI3	LIB3 deflated by annual percentage change in US wholesale price index	%
G7FB	Fiscal balance/GNP, average for the group of 7 major industrial countries	%
PM	Import unit values in US$: annual percentage change	%
PX	Export unit values in US$: annual percentage change,	%

Sources
IMF (1981, 1986b, 1989b)
IMF, International financial statistics (annual)

Each of these variables takes only eight different values within the sample, one for each year:- the value is the same for all countries in a given year. (In the case of PM, there are two values for each

year: one for fuel exporters and one for other countries. The same is true for PX). In a sense these are partitioning variables, but they are ratio-scaled. The rationale for including them is akin to the reasons for including categorical variables in regression equations - indeed they can be transformed into categorical variables - but by being ratio-scaled, they contain more information than nominal or ordinal data.

IV.4 Categorical variables

The final additions to the data set were nine categorical variables, as set out in Table IV.5. Variables HID, KMX, LIC, OIL, PX and REG are based on IMF definitions (IMF, 1989b).

Table IV.5: Categorical variables

ARS	**Debt-servicing status: see chapter III.**							
	Coding:					0	1	2 3 4 9
DATE	**Date: Year:**						1979	to 1986
HID	**Heavily indebted country:**	Yes:	1	No:	0			
ID	**Country coding:**	Sequence:					1	to 55
KMX	**Capital imports or exports:**							
	Capital exporter:	1	Market borrower:		2			
	Official borrower:	3	Diversified borrower:		4			
LIC	**Low income country:**	Yes:	1	No:	0			
OIL	**Status as oil exporter:**							
	Oil exporting countries:				1			
	Net oil exporters:	2	Net oil importers:		3			
PRX	**Predominant export:**							
	Fuel:	1	Minerals:	2	Agricultural:		3	
	Manufacturing:	4	Service + remittance country:		5			
REG	**Geographical region:**							
	Asia:	1	Europe:	2	Middle East:		3	
	Western Hemisphere:	4	Africa:		5			

IV.5 Statistical properties of the data

IV.5.1 Introduction

Here we are concerned with the primary data set of 70 variables. Each of the eight external (environmental) variables has only eight distinct values in the data set (one for each year), and they may be regarded as, in effect, qualitative variables.

The statistical properties of the models that are described in later chapters depend, inter alia, on the distributional properties of the population from which the data sample was drawn. Moreover, one of the techniques that is utilised, linear discriminant analysis, is based on the assumption that the discriminating variables have a multivariate normal distribution. For these reasons, it is necessary to explore the distributional properties of the sample data, in order to draw inferences about the distribution of the parent population. In this chapter, this exploration will be limited to the marginal distribution of each variable. Insofar as multivariate normality is concerned, the shape of the marginal distribution provides a one-way test: marginal normality being a necessary but not sufficient condition for multivariate normality.

Tests for multivariate normality are available, but it will be appropriate to confine such testing to the reduced data set of variables that will actually be used in the statistical analysis. This matter will be taken up in chapter VII.

In undertaking this analysis, two approaches were used, and the results are reported in sections IV.5.2 and IV.5.3 respectively. First, an exploratory analysis was made through visual inspections of the histogram and the normal probability plot of the sample data for each variable. Secondly, three separate tests for normality were applied: the non-parametric Kolmogorov-Smirnov test (Conover, 1980), secondly a standard chi-square test, and finally the Jarque-Bera test, which is based on the sample coefficients of skewness and kurtosis (Jarque and Bera, 1980). At each stage, separate analyses were undertaken for the arrears and non-arrears sub-samples.

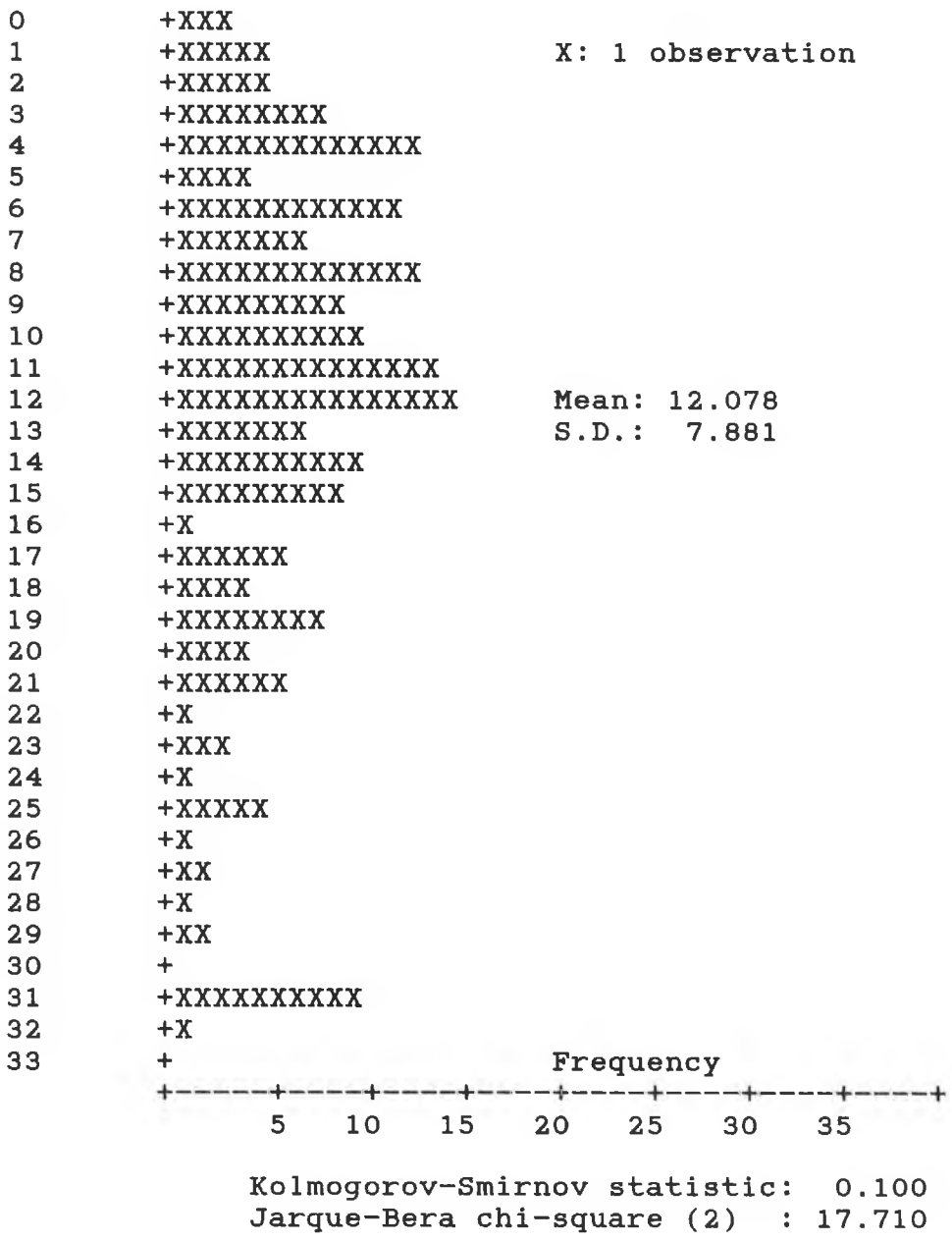
IV.5.2 Graphical analysis and formal tests

The histograms and normal plots were generated by the BMDP programme 5D, the Kolmogorov-Smirnov test utilised the SPSS-X programme NPAR TESTS, and the Jarque-Bera and standard chi-square tests were

obtained using the SHAZAM programme OLS.

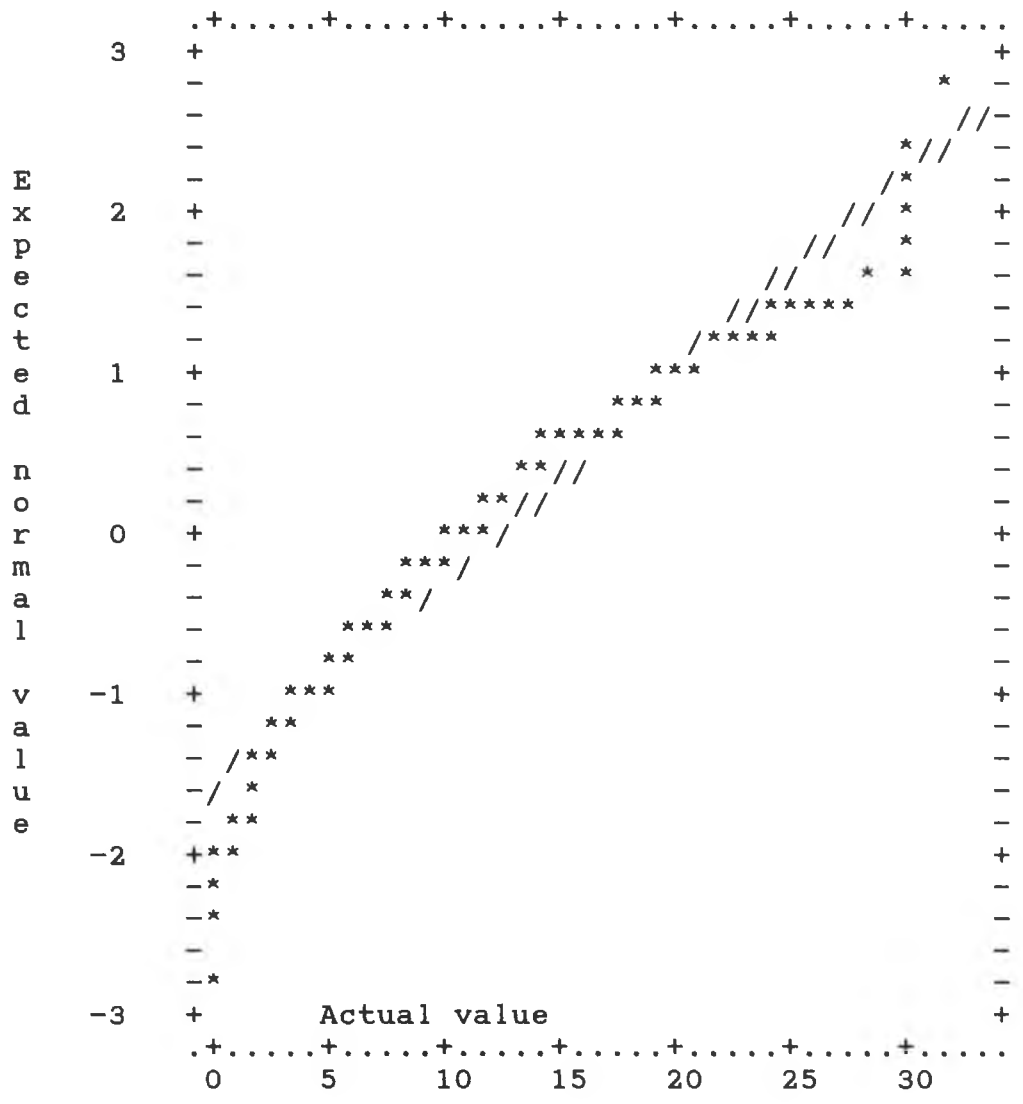
The sample histogram will give an indication of the overall shape of the data. The normal probability plots map the sample data against theoretical normal deviates, and will approach linearity if the data are normal. Figures IV.1 - IV.3 show the histograms and plots for a subset of the data: those for INPS indicate normality; those for DCPI indicate right skewness; those for CARAG indicate more probability in the tails of the distribution than occurs for normally distributed variables. At the five per cent significance level, the critical value of chi-square (2 d.f.) for the Jarque-Bera test is 5.991, while the critical value of the Kolmogorov-Smirnov statistic for a two-tailed test is 0.096.

Figure IV.1(a): Histogram of variable DCPI
200 non-arrears cases



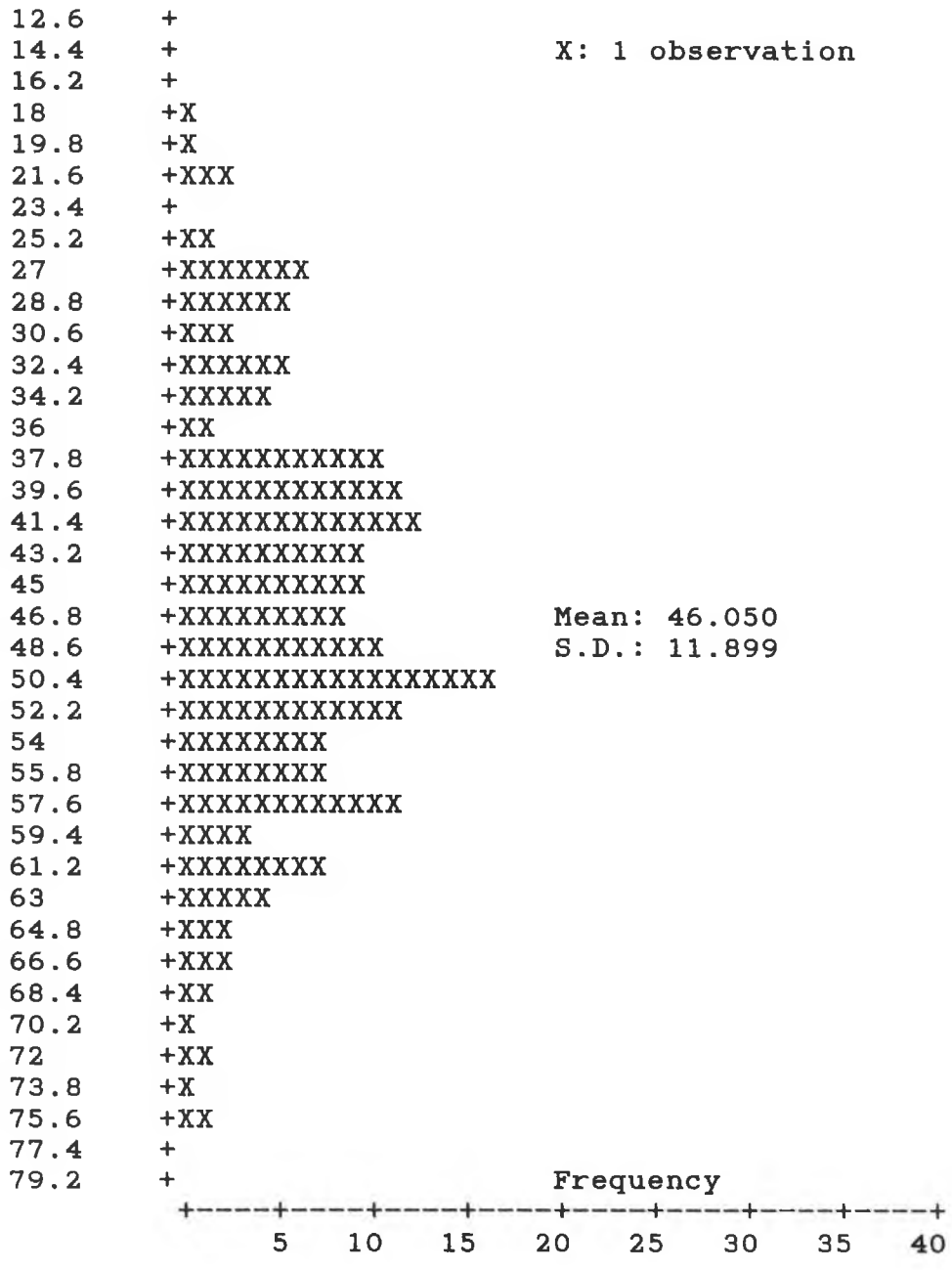
Kolmogorov-Smirnov statistic: 0.100
Jarque-Bera chi-square (2) : 17.710

Figure IV.1(b): Normal plot of variable DCPI
200 non-arrears cases



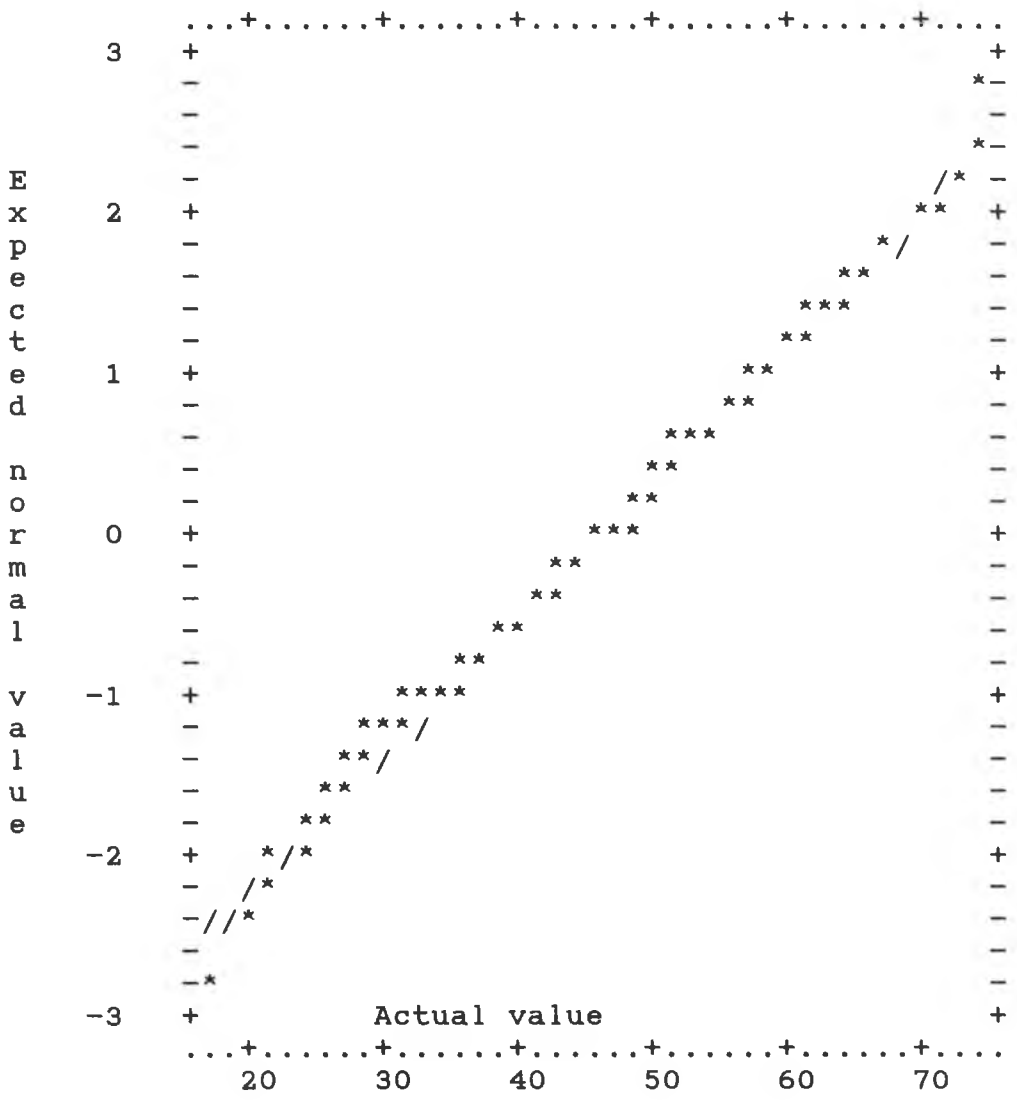
Values from normal distribution would lie on the line indicated by the symbol /

Figure IV.2(a): Histogram of variable INPS
200 non-arrears cases



Kolmogorov-Smirnov statistic: 0.030
Jarque-Bera chi-square (2) : 0.983

Figure IV.2(b): Normal plot of variable INPS
200 non-arrears cases



Values from normal distribution would lie on the line indicated by the symbol /

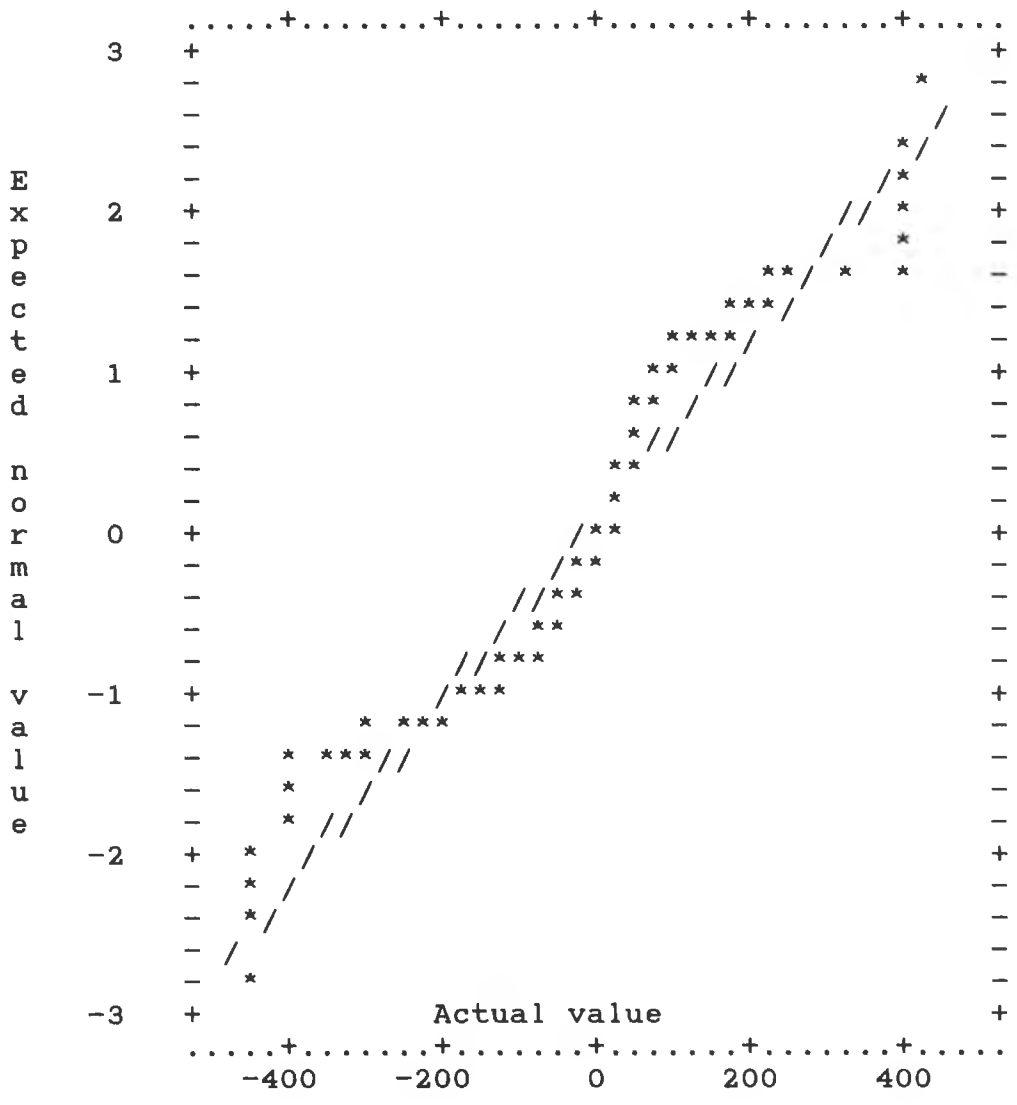
Figure IV.3(a): Histogram of variable CARAG
200 non-arrears cases

```
Interval

-475    +
-450    +X                    X: 1 observation
-425    +XXXXX
-400    +XXXXXXXXX
-375    +X
-350    +X
-325    +
-300    +XX
-275    +X
-250    +
-225    +XX
-200    +XXXX
-175    +XX
-150    +XXXXX
-125    +XXXXXX
-100    +XXXXX
-75     +XXXXX
-50     +XXXXXXXXX
-25     +XXXXXXXXXXXXXXXX
0       +XXXXXXXXXXXXX
25      +XXXXXXXXXXXXXXXXXXXXXXXXXXXXX
50      +XXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
75      +XXXXXXXXXXXXX
100     +XXXXXXXXXX           Mean: -15.861
125     +XXXXXXXX             S.D.: 178.829
150     +X
175     +XXX
200     +XX
225     +X
250     +XX
275     +X
300     +
325     +
350     +X
375     +
400     +XX
425     +XXXXXXXX
450     +                     Frequency
        +----+----+----+----+----+----+----+----+
             5    10   15   20   25   30   35   40

        Kolmogorov-Smirnov statistic:  0.149
        Jarque-Bera chi-square (2)   : 15.307
```

Figure IV.3(b): Normal plot of variable CARAG
200 non-arrears cases



Values from normal distribution would lie on the line indicated by the symbol /

Overall, 45 out of 70 variables had near-linear normal probability plots, while the plots and histograms of the remaining 25 were characterised by varying combinations of right-skewness and a heavier weight in the tails than the normal distribution. Skewness was the predominant aspect of departures from normality, and no variables had a symmetric heavy-tailed sample distribution in each sub-sample.

This complex picture is not surprising. Some of the variables measure country size: for example, LABF, POPN, and RY (real GDP). A priori, it is clear that the sample contains a bigger proportion of small countries than large, and therefore these variables will have skewed distributions. All of the other obviously skewed variables are ratios (a category that includes growth rates), which again is explicable: skewness can arise from taking the ratio of two normal variates (Barnes, 1982), and moreover ratio distributions in general tend to be skewed (Frecka and Hopwood, 1983).

A second consequence of the ratio origins of many of the variables is that the tails of their distributions may be heavier than those of the normal distribution. For example, in the case of a sum/sum ratio, it is quite feasible to have a heavy-tailed distribution if both numerator and denominator have

significant concentrations of values near zero.

IV.5.3 Tests of univariate normality: conclusions

For the formal statistical tests, the null hypothesis that the parent population is normally distributed was tested at a significance level of five percent. The results of the three tests are not consistent with each other. Notably, in the case of 15 variables in the arrears sub-sample, and 18 variables in the non-arrears sub-sample, the chi-square test rejects the null hypothesis, whereas the other two tests do not reject it, and moreover the normal probability plots are close to linearity in these cases. The chi-square test is based on an arbitrary arrangement of the sample into a small number of class intervals, and for this reason is likely to be less reliable than the other tests (Kendall and Stuart, 1979: vol. 2, p.466).

If we ignore the chi-square test, the picture is a little clearer: on the basis of both the Kolmogorov-Smirnov and Jarque-Bera tests, the null hypothesis is not rejected for 20 variables on both sub-samples, nor for a further 14 variables on the arrears sub-sample only, nor for 2 other variables on the non-arrears sub sample only. Moreover, in each

of these cases the normal probability plots are close to linearity. However, for the remaining variables, the results are mixed.

With two formal tests and the informal visual inspection of the normal plot to choose from, each possible pattern of rejection versus non-rejection is in fact adhered to by at least one variable, although the major categories are: rejected only by Jarque-Bera, rejected only by Jarque-Bera and visual inspection, rejected on all three approaches. Furthermore, the pattern is not consistent across the two sub-samples. The conclusion from this complex picture is that the hypothesis of joint normality must be rejected for the data set as a whole. However, in chapter VII, selected subsets of the data set will be submitted to a test for multivariate normality.

The results reported above indicate directions for further research. First, there is a discordance between the results of the formal tests, and this merits investigation. Secondly, further research is required to establish the distributional properties of those variables that are not univariate normal.

IV.6 Transformation to normality; treatment of outliers

IV.6.1 Analysis of transformations

In this section the possibility is investigated of finding transformations of the data set that would yield a set of marginally normally distributed variables.

McLeay (1987) observes that ratios and other variables may be divided into three groups:

(i) Ratios of sums, which are bounded by zero;

(ii) Ratios of differences to sums, or of sums to differences;

(iii) Ratios of differences which, like type (ii), may take any real value.

Of the 70 variables introduced above, 38 fall into category (i), 27 into category (ii) and 5 into category (iii). According to McLeay, category (iii) may generally be treated like category (ii).

Because category (i) variables are bounded by zero, it is possible to apply the Box-Cox procedure (Box and Cox, 1964; Johnson and Wichern, 1982; Emerson and Soto, 1983). This involves using maximum likelihood methods to find the power transform that most closely

approximates to normality.

The Box-Cox procedure cannot in general be applied directly to variables in categories (ii) or (iii), because the likelihood function is not defined for negative arguments. One approach is to shift the sample values to the right by adding an appropriately chosen constant (Johnson and Wichern, 1982: p.161). In an exploration of the present data set, a lower bound of at least unity was imposed on each variable in the sample. If the sample minimum was at least unity initially, no adjustment was made; otherwise [1.0-(sample minimum)] was added to each sample value. The lower bound was set at unity rather than zero in order to avoid the numerical difficulties associated with transformation of very small numbers, in particular logarithmic transformation. All the variables of types (ii) and (iii) were re-scaled in this way, as also were 16 of 38 type (i) variables.

The Box-Cox transformation is y=[(x**z)-1.0]/z for z non-zero and y=log(x) otherwise, where x is the untransformed variable. [The exponential operator is denoted by **.] On the hypothesis that y is normally distributed for some value of z, the procedure is to find the value of z that maximises the probability that the given sample values would be observed. In the present case, for each variable, the likelihood

function was maximised over the interval [-3,3]. The maximum was located by a search procedure: a grid of values in the interval at spacings of 0.1 was established, and the value of the likelihood function was computed at each point. Inspection of the graph of the likelihood function confirmed, in each case, that the grid was sufficiently fine for the computed maximum to be a good approximation to the true maximum. Having undertaken this analysis for the complete sample of 440 cases, it was then undertaken separately for the 'arrears' and 'non-arrears' sub-samples. Examination of the graphs of the three likelihood functions showed that where there were differences in the maximising values of z (and the differences were always small), the likelihood functions always had such flat maxima that the differences between the maximising values of z were of little significance. In several cases there was more than one local maximum on the interval [-3,3], and in such cases the global maximum was identified by inspection. The results of this analysis are set out in Table IV.6.

Table IV.6: Optimal power transforms

Maximising value of z	Variables	No.
-3.0 < z <-0.375	-	none
-0.375 < z < 0.0	LABF MCOV RY	3
z = 0.0 (log transform)	EFMT POPN RTD STPD	4
0.0 < z < 0.5	AGRP GDPX INPX INPY RDPC RYPC TDPX TDSR STPY XYRA *DCPI +DCPIG +STPDG	13
z = 0.5 (square root)	EFIR EXIM TDPY TSPD TSPX TSPY XDPD XPM12 *DMN22 *ILMAG *MCOVG *RTDG	12
0.5 < z < 1.0	INDP INVR MLPDG OCTD PDSPX SIRA XPP12 *DDCR *DMN2 *LABFG *NBTTG *TDPYG *TSPDG *TSPXG *TSPYG	15
z = 1.0 (no transform)	COPC DCPY *DFIN *DGDP *DPOP *FRQPY *MGRO *RYPCG *TDPXG *TDRA *XGRO +BHPCA +CARAG	13
1.0 < z < 1.5	INPS MYRA NBTT SERP *CARA *DCPR *DIND *NARY *PSBR +PSBRG	10
1.5 < z < 3.0	-	none

Notes:
*: category (ii) +: category (iii)
All other variables are of category (i).

In Table IV.6, category (ii) variables are marked *, category (iii) variables are marked +, and all other

variables are in category (i). For only three variables does the best transform have a negative exponent, while for all the others the maximising exponent falls between zero (i.e. log transform) and 1.5. In only 13 cases is it optimal to make no transformation (i.e. z=1).

The use of an arbitrary constant to avoid negative sample values is unsatisfactory. First, as McLeay points out (1987: p.163) it involves some loss of information. Secondly, there is no reason to suppose that the population values are bounded below by the sample minimum. An alternative approach, resting on the results reported by McLeay, would be to assume that ratios of types (ii) and (iii) have the t-distribution. In each case, three parameters would have to be estimated, representing degrees of freedom, location, and scale. Having done that, a transformation to normality could be achieved through utilisation of the relationship between the t, chi-square and standard normal distributions.

This alternative has not been pursued, and most of the statistical analyses that are reported in later sections of this thesis are based on the untransformed, but Winsorized, data set. The only exceptions are the principal components and discriminant analyses: although the central

investigations concern untransformed data, the transformed data were explored briefly, with results that will be reported in due course.

IV.6.2 Transformation versus Winsorization

The decision to work with untransformed data is based on the following considerations. First, previous research indicates that, in practice, linear discriminant analysis is robust to departure from normality. According to Taffler:

> Multivariate normality is an important assumption [i.e. underlying linear discriminant analysis] although in sample-based research it is the presence of extreme observations rather than the general shape of the multivariate distribution that is the key. (Taffler, 1986: p.4.)

This motivates the use of the Winsorizing procedure, to be described below.

Secondly, other statistical techniques will be used whose validity does not depend on the assumption that the data are multivariate normal. Given these two points, and in view of the objective of this thesis, it was decided not to proceed further at this stage with analysis of the distributional properties of the sample data. McLeay's research is concerned with microeconomic data, and there is no reason to suppose

that his conclusions concerning t-distributions would carry over to the macro data used here. In particular, the previous section suggests that the principal dimension of non-normality in the sample is skewness, rather than symmetry with weight in the tails. It would be possible to look for candidates other than the t-distribution, but this would involve a major detour from the central theme. Moreover, because it is not the case that every variable has the same type of distribution, there would remain the problem of choosing the correct statistical tests of significance.

As for the normal transforms described above, there are three reasons for rejecting their use. First, there is the problem of the arbitrary parameter of location. Secondly in the principle components analysis of chapter V, components are extracted from the untransformed data set that have a ready interpretability: similar ratios turn out to be highly loaded on the same component. This property does not hold when the transformed data are used. Thirdly, according to Eisenbeis (1977), transformations to normality can change the relationships among the variables.

The data set that is analysed in later chapters has been Winsorized: that is to say, observations more

than 2.5 standard deviations from the sample mean have been brought in towards the centre of the distribution (Barnett and Lewis, 1984). This has been carried out using a recursive technique:

(i) for each variable, the sample mean M1 and standard deviation S1 were computed;

(ii) then the trimmed mean M1' and standard deviation S1' were computed for the censored sub-sample remaining after deletion of observations outside the limits 2.5S1 either side of M1;

(iii) these outliers were then assigned values M1'-2.5S1' or M1'+2.5S1', according as they lay below or above the mean;

(iv) the mean M2 and standard deviation S2 were computed for the reconstructed full size sample, and the procedure continued for another iteration;

(v) the procedure was stopped when there were no outliers at stage (ii).

The arrears and non-arrears sub-samples were Winsorized separately. In most cases, convergence was achieved in no more than four rounds. For the arrears sub-sample, three variables took five rounds to converge and one took seven; for the non-arrears sub-sample, one variable took six rounds and two took seven.

Details of the initial number of outliers for each variable are shown in Table IV.7. The tails of a normal distribution outside 2.5 standard deviations either side of the mean contain 1.24 per cent of the total probability. On the hypothesis that the data

are normally distributed, the expected number of outliers is therefore between one and two for the arrears sub sample (because 1.24 per cent of the sub-sample size (102) is 1.26), and between two and three for the non-arrears cases. Table IV.7 reveals that in the arrears sub-sample, only 39 out of 85 variables have one or two outliers. The evidence of non-normality is more pronounced in the non-arrears data, where only 15 variables have the expected two or three outliers; 54 have more than three and not more than eight, while seven have nine or more - including one variable with fifteen outliers.

Table IV.7: Initial numbers of outliers

Arrears sub-sample		Non arrears sub-sample	
Numbers of Outliers	Variables	Numbers of Outliers	Variables
0	6	0 or 1	9
{1 or 2}	39	{2 or 3}	15
3 or 4	32	4,5,6,7,8	54
5,6,7	8	7 to 15	7

Notes:
{} indicates expected range, given normality.

Outliers are taken to be cases outside 2.5 standard deviations either side of the mean.

The purpose of Winsorization is threefold. First, the presence of outliers can increase sample variance and decrease precision (Cochran, 1963). Secondly, the values of parameter estimates can be dominated by

the values of the outliers - a problem which, if left uncorrected, is a counterpart of the tendency of human analysts to overlook 'regression towards the mean'. Thirdly, the simulation studies of Lachenbruch et al (1973) indicate that linear discriminant analysis suffers least from non-normality when the discriminating variables are bounded above and below.

Appendix IV

The variables that are discussed in Chapter IV may be classified and grouped as follows:

Policy indicators

DCPI Consumer price inflation
DCPIG Growth of DCPI
DDCR Growth of domestic credit
DMN2 Growth of money supply
DMN22 DMN2, lagged two years
PSBR Public sector deficit/GDP
PSBRG Growth of PSBR

Debt

COPC New loan commitments/population
GDPX Public external medium and long term debt/exports
NARY (International reserves minus total external debt)/GDP
RDPC Real total external debt/population
RTDG Growth of real total external debt
TDPX Total external debt/exports
TDPXG Growth of TDPX
TDPY Total external debt/GDP
TDPYG Growth of TDPY

Debt service

INPX Interest on total external debt/exports
INPY Interest on total external debt/GDP
PDSPX One-year ahead projection of debt service/current exports
TDSR (Total external debt service plus total external short term debt)/exports
TSPX Total external debt service/exports
TSPXG Growth of TSPX
TSPY Total external debt service/GDP
TSPYG Growth of TSPY

Short-term debt

STPD Total external short-term debt/total external debt
STPDG Growth of STPD
STPY Total external short-term debt/GDP

Characteristics of debt and debt-service

EFIR Effective interest rate
EFMT Effective maturity
INPS Interest/debt service
MLDPD Debt outstanding to multilateral institutions/total medium and long term external debt
OCTD Official creditors/total external debt
TSPD Total external debt service/total external debt
TSPDG Growth of TSPD

Balance of payments
BHPCA Balancing item on capital account/balance on current account
CARA Current account balance/GDP
CARAG Growth of CARA
EXIM Exports/imports
FRQPY (Current account balance minus principal repayments on total external debt)/GDP

International reserves
MCOV Import cover
MCOVG Growth of MCOV
ILMAG Growth of international reserves

Strength of trading sector
MGRO Growth of real imports
NBTT Terms of trade
NBTTG Growth of NBTT
TDRA External trade balance/GDP
XDPD Index of price competitiveness
XGRO Growth of real exports
XPM12 Exports to two largest markets/exports
XPP12 Exports of first plus second principal export products/exports

Openess of economy
MYRA Imports/GDP
XYRA Exports/GDP

Structural indicators
AGRP Agricultural output/GDP
DCPY Domestic credit/GDP
DPOP Growth of population
INDP Industrial output/GDP
INVR Real fixed investment/GDP
LABFG Growth of labour force
RYPC Real GDP/population
SERP Output of services sector/GDP
SIRA Savings/investment

Size of economy
LABF Labour force
POPN Population
RTD Real total external debt
RY Real GDP

Economic growth
DFIN Growth of real fixed investment
DGDP Growth of real GDP
DIND Growth of real industrial production
DCPR Growth of real private consumption
RYPCG Growth of (real GDP/population)

External environment and creditor behaviour

WOP	Annual growth rate of world output
ICY	Annual growth rate of real GNP: industrial countries
LIB3	London interbank offered rate on US$ three-month deposits
RLI3	LIB3 deflated by annual percentage change in US wholesale price index
G7FB	Fiscal balance/GNP, average for the group of 7 major industrial countries
PM	Import unit values in US$: annual percentage change
PX	Export unit values in US$: annual percentage change,

V Preliminary analysis of data set

V.1 Introduction

In this chapter, the structure of the data set is explored. This is done in two different ways, each for a particular purpose.

First, it is necessary to identify the significant dimensions of the data set, and to group the variables in terms of their loadings on those dimensions. In section V.2, this is achieved through a principal components analysis, a statistical technique that enables a large data set to be reduced to a small, parsimonious set of key variables. Compared with informal methods of achieving parsimony, the technique involves less risk of dropping information. A related objective is to identify the existence and location of multicollinearity within the data set.

Principal components analysis has been applied for these purposes in earlier work on country risk: for example Taffler and Abassi (1984) and Somerville (1985). Extracting components in itself is not directly applicable to early warning; Dhonte (1975) has explored this possibility, and he reports a

discouraging type I error rate of over 30 per cent.

A second possible use for the principal components analysis would be to use the derived component scores as input variables for the early warning models, instead of using the original variables. For example, in a study of corporate bankruptcy, Mensah (1984) used factor scores as input variables to a logit analysis. However, in the interests of easy interpretability, Mensah's precedent will not be followed.

In the second part of this chapter, the extent of homogeneity of the 55 countries within the data set is examined. These countries are highly diverse. Of course, some of this diversity is picked up by continuous variables in the data set: for example, size, economic structure, real income levels. In other respects, the diversity is reflected by categorical measures. To the extent that this within-sample diversity is not reflected by the sample variables, it is inappropriate to treat the sample as homogeneous and to estimate statistical relationships across it without any form of partitioning. Therefore, it is necessary to explore the data for evidence of natural partitions, and hierarchical cluster analysis is used for this purpose in section V.3.

V.2 The dimensions of the data set:

V.2.1 Choice of statistical technique

The first matter to settle concerns the technique to be used: the choices here include principal components analysis, and the set of techniques known collectively as factor analysis. All of these techniques may be used to reduce a set of variables to a smaller set of composites. According to Harman (1976: p.15)

> An important property of [principal components analysis] is that each component in turn makes a maximum contribution to the sum of the variances of the n variables...In contrast...the traditional, or classical factor analysis model is designed to maximally reproduce the correlations [among the variables].

while Green (1978: p.403) writes that

> in the author's view the principal components model has much to offer as a basic factor model.....Unless the researcher has very good reason to believe that communalities differ markedly across variables, it is probably a prudent approach to use the components model, followed by Varimax rotation or some other such orthogonal transformation procedure.

In this thesis, factor analytic procedures are utilized as methods of data reduction, rather than to explain or reproduce the correlations between the variables. Consequently, and given the views cited above, principal components analysis is the primary

approach used here.

However, other factor analytic methods were applied to the data set: these included principal axis factoring, alpha factoring, image factoring, unweighted least squares, generalised least squares, and maximum likelihood. The factors that were extracted by these methods were, after varimax rotation, virtually identical with the principal components: both in terms of the identities of the highly loaded variables, and of the magnitudes of the factor loadings.

References for these statistical techniques are Green (1978) for principal components analysis, and Harman (1976) for the various methods of factor analysis. The processing was carried out using the FACTOR programme of the SPSS-X software package.

V.2.2 The principal components analysis

A principal components analysis was carried out on the primary data set of 70 variables, using 440 observations: i.e., including arrears, non-arrears and weak-year cases. The number of components to be extracted was determined by the criterion that the associated eigenvalues should exceed unity (Kaiser, 1959 cited in Green, 1978: p.364). On this basis, 19 components are retained, representing 78.9 per cent of the total variance within the data set. After varimax rotation, these components may be reified by inspection to yield the labelling that is shown in Table V.1, and summarised as follows (eigenvalues in parentheses):

#1	Debt or debt service ratios	(8.7)
#2	Inflation	(7.7)
#3	Debt profile	(6.1)
#4	Economic structure	(4.8)
#5	Economic growth	(3.8)
#6	Current account	(2.9)
#7	Size	(2.6)
#8	Trade balance, diversification	(2.4)
#9	Growth of debt service	(2.1)
#10	Growth of reserves	(1.9)
#11	Growth of debt	(1.8)
#12	--	(1.6)
#13	Short term debt	(1.5)
#14	--	(1.5)
#15	Terms of trade	(1.3)
#16	Share of service sector	(1.2)
#17	--	(1.1)
#18	Public sector deficit	(1.1)
#19	--	(1.0)

Table V.1: Principal components analysis

No. and loading	COMPONENT NAME (Incremental and cumulative % of variance)
# 1	DEBT OR DEBT SERVICE/EXPORTS OR GDP (+12.4%, 12.4%)
0.86	TSPY (Total debt service/GDP)
0.85	INPY (Total interest/GDP)
-0.76	NARY (Net assets/GDP)
0.76	TDPY (Total debt/GDP)
0.64	STPY (Short term debt/GDP)
0.55	RDPC (Real debt per capita)
0.53	TSPX (Total debt service/exports)
0.52	INPX (Total interest/exports)
0.45	TDSR (Total debt service ratio)
0.40	PDSPX (Projected debt service ratio)
0.42	[XYRA (Exports/GDP)]
# 2	INFLATION (+10.9%, 23.3%)
0.85	DCPI (Consumer prices: rate of increase)
0.83	DMN2 (Monetary growth rate)
0.81	DDCR (G.R., domestic credit)
0.68	DMN22 (DMN2, lagged two years)
0.63	[INPX (Total interest/exports)]
0.59	[TSPX (Total debt service/exports)]
0.53	[DSPX (Projected debt service ratio)]
0.51	[TDSR (Total debt-service ratio)]
0.50	[TDPX (Total debt/exports)]
-0.45	[XYRA (Exports/GDP)]
0.41	[GDPX (Public debt/exports)]
# 3	DEBT PROFILE: MATURITY, INTEREST RATE (+8.8%, 32.1%)
-0.89	EFMT (Effective maturity)
0.85	TSPD (Total debt service/debt)
0.64	EFIR (Effective interest rate
-0.48	OCTD (Official creditors/total debt)
-0.44	INPS (Interest payments/debt service)
-0.66	[GDPX (Public sector debt/exports)]
-0.56	[TDPX (Total debt/exports)]
0.51	[NARY (Net assets/GDP)]
-0.49	[TDPY (Total debt/GDP)]
0.45	[RYPC (Real GDP per capita)]
-0.41	[PDSPX (Projected public debt serv./exports)]

G.R. denotes annual percentage growth rate

.../

.../

# 4	ECONOMIC STRUCTURE
	(+6.8%, 38.9%)
-0.79	AGRP (Agriculture/GDP)
-0.76	MLDPD (Multilateral debt/medium+long debt)
0.76	INDP (Industry/GDP)
0.65	RYPC (Real GDP per capita)
0.58	RDPC (Real debt per capita)
-0.53	OCTD (Official creditors/total debt)
0.49	INVR (Fixed investment/GDP)
# 5	ECONOMIC GROWTH
	(+5.4%, 44.4%)
0.86	DGDP (G.R., real GDP)
0.77	DIND (G.R., industrial production)
0.75	DCPR (G.R., private consumption)
0.72	DFIN (G.R., real fixed investment)
0.58	MGRO (G.R., imports)
# 6	CURRENT ACCOUNT
	(+4.2%, 48.5)
0.89	CARA (Current account balance/GDP)
0.89	FRQPY (Financing requirement/GDP)
0.77	COPC (New loan commitments per capita)
0.75	SIRA (Savings/investment)
0.44	EXIM (Exports/imports)
0.41	CARAG (G.R., financing requirement/GDP)
# 7	SIZE
	(+3.7%, 52.3%)
0.91	POPN (Population)
0.90	LABF (Labour force)
0.81	RY (Real GDP)
0.77	RTD (Real total debt)
# 8	TRADE BALANCE, DIVERSIFICATION
	(+3.5%, 55.8%)
0.68	TDRA (Trade balance/GDP)
0.61	XPM12 (Share of first two export markets)
0.58	EXIM (Exports/imports)
0.41	XPP12 (Share of first two exp't products)
# 9	GROWTH OF DEBT SERVICE
	(+3.1%, 58.8%)
0.92	TSPDG (G.R., total debt service/debt)
0.90	TSPXG (G.R., total debt service/exports)
0.87	TSPYG (G.R., total debt service/GDP)
# 10	GROWTH OF RESERVES
	(+2.7%, 61.5%)
0.93	ILMAG (G.R., international reserves)
0.91	MCOVG (G.R., import cover)

.../

```
 .../
-----------------------------------------------
 # 11¦ GROWTH OF DEBT
     ¦ (+2.6%, 64.0%)
 0.86¦ RTDG (G.R., real total debt)
 0.73¦TDPXG (G.R., total debt/exports)
 0.54¦TDPYG (G.R., total debt/GDP)
-----------------------------------------------
 # 12¦
     ¦ (+2.3%, 66.3%)
 0.80¦RYPCG (G.R., real GDP per capita)
-0.73¦TDPYG (G.R., total debt/GDP)
 0.51¦ XDPD (Index of price competitiveness)
-----------------------------------------------
 # 13¦ SHORT TERM DEBT
     ¦ (+2.2%, 68.5%)
 0.73¦ STPD (Short term debt/total debt)
 0.54¦ STPY (Short term debt/GDP)
 0.51¦ TDSR ([Debt serv.+short debt]/exports)
-----------------------------------------------
 # 14¦
     ¦ (+2.1%, 70.6%)
 0.65¦ DPOP (G.R., population)
 0.59¦XPP12 (Share of first two exp't products)
-0.46¦ INPS (Interest payments/debt service)
-----------------------------------------------
 # 15¦ TERMS OF TRADE
     ¦ (+1.9%, 72.5%)
 0.79¦NBTTG (G.R., terms of trade)
 0.69¦ NBTT (Terms of trade)
-----------------------------------------------
 # 16¦ SHARE OF SERVICE SECTOR
     ¦ (+1.7%, 74.2%)
 0.76¦ SERP (Services/GDP)
- - - - - - - - - - - - - - - - - - - - - - - -
-0.41¦[DCPIG (G.R. of consumer price inflation)]
-----------------------------------------------
 # 17¦
     ¦ (+1.6%, 75.8%)
 0.72¦ MYRA (Imports/GDP)
 0.53¦BHPCA (Balancing item/C.A. balance)
-0.41¦STPDG (G.R., short term debt/debt)
-----------------------------------------------
 # 18¦ PUBLIC SECTOR DEFICIT
     ¦ (+1.5%, 77.4%)
 0.74¦PSBRG (G.R., public sector deficit/GDP)
 0.54¦ PSBR (Public sector deficit/GDP)
-----------------------------------------------
 # 19¦
     ¦ (+1.5%, 78.9%)
 0.65¦ XGRO (G.R., exports)
-0.60¦LABFG (G.R., labour force)
-----------------------------------------------
```

Where the component loading is below 0.4, the variable has been suppressed from Table V.1, while square brackets indicate variables that, while loaded on a given component, do not conform to the general interpretation of that component. Four components (numbers 12, 14, 17 and 19) have, in each case, diverse variables with high loadings. Consequently, no interpretation is attempted for these cases.

Only two variables do not appear in Table V.1. MCOV (import cover) is absent, but its growth rate is present. Secondly, DCPY (domestic credit/GDP) is absent, but the growth rates of domestic credit and of the money supply (both unlagged and lagged) are present.

In similar research on corporate data, the number of components retained on Kaiser's criterion has tended to be smaller. At the same time, the proportion of variance explained by retained components has tended to be greater than that reported here. For example Taffler and Sudarsanam (1980) extracted nine principal components from a set of 80 corporate variables for a cross section of companies in the financial year 1976-1977. All had eigenvalues exceeding unity, and the proportion of total variance explained was 91.3 per cent. The first six components explained 85.7 per cent. Ezzamel et al

(1987) used a set of 53 ratios for 1973, 1977 and 1981. The numbers of factors retained in each year were respectively 11, 15 and 10, and the proportions of the variance that these factors explained were respectively 79 per cent, 81 per cent and 86 per cent.

There are at least two possible explanations for the rather large number of retained components in the research reported here. First, country data is likely to contain more natural dimensions than corporate data. Some country ratios have corporate counterparts: for example, some of the ratios involving debt and debt-service. For others, such as indicators of inflation, population and labour force, and sectoral shares, there is no counterpart in the data that are analysed in the accounting literature. Secondly, the corporate studies typically use a pure cross-section, while in this thesis panel data are used.

Taffler and Abassi (1984) retained 10 principal components in an analysis of panel data for LDCs. These ten explained 84 per cent of the total variance in the data set; however, their data included only 42 variables, compared with 70 in this thesis. Notably, of the 23 variables that are shown in Table V.1 as being highly loaded on components 11 to 19, only two

(imports/GDP and growth rate of exports) were used by Taffler and Abassi. Consequently, their results are not directly comparable with those reported here.

Otherwise, the results of this analysis are broadly satisfactory. The first seven components account for more than half the total variance within the data set, and they represent the key elements of debt and debt service (both level and structure), inflation, economic structure, economic growth, the current account, and the size of the economy. They also cover seven of the eight dimensions that were identified in chapter IV as having been used in earlier work on early warning, and three of the five categories of Suttle (1989). In both cases, the missing categories consist of variables reflecting the external environment (see Table IV.4): these are not included in the principal components analysis.

The principal instances of variables being loaded on the 'wrong' component arise with the collection of ratios of debt and debt-service to GDP and exports. These are highly loaded on components 2 and 3, as well as on the component (#1) with which they are primarily identified. Earlier studies have yielded similar mixing within components: for example in Taffler and Abassi (1984), inflation is highly loaded on a component which is reified as 'external trade'.

The results of this analysis are used in later chapters as an aid to decisions about the selection of variables. Where the statistical methods are based on stepwise selection of variables, then an interactive procedure is used. The results presented here are used, in conjunction with a correlation analysis, to determine the route to be taken at each step.

V.2.3 Transformations to normality

The principle components analysis was applied to the data set after subjecting each variable to the optimal power transform as described in chapter IV.6. The components thus extracted, after Varimax rotation, had no ready interpretability in terms of the identities of the variables that were highly loaded on them. This point is referred to in chapter IV, along with other considerations upon which is based the decision to work with untransformed, but Winsorized, variables.

V.3 Partitions within the data set: hierarchical cluster analysis

V.3.1 Introduction

The objective of factor analysis, including principal components analysis, is to reduce the number of variables in a data set to a smaller number of composites. In contrast, the purpose of hierarchical cluster analysis is to reduce the number of cases in the data set - i.e. to group them. Cluster analysis, which is based on proximity of cases in the space of variables, is described in Green (1978), Everitt (1980) and Gordon (1987).

In this chapter, cluster analysis is used to search for any natural partitions that may exist within the set of continuous explanatory variables (i.e. the primary data set of 70 financial and economic variables). A number of agglomerative methods are available for hierarchical clustering, and unfortunately there are no clear theoretical grounds for making a choice of method. Everitt (1980: pp. 100-101) surveys a number of empirical studies that compare the results of applying different methods to given data sets. The methods that perform best in those studies are Ward's method, weighted average

between groups, and complete linkage (furthest neighbour), and these are therefore used in this thesis. Clustering was carried out by the CLUSTER programme of the SPSS-X software package.

V.3.2 Choice of variables for clustering

Three different selections of variables were made for the cluster analysis. First, variables representative of those used in other applied work were identified. The material cited in chapter IV.1 was consulted, and the variables listed in Table V.2 were selected.

Table V.2: Clustering variables: first selection

CARA (current A/C balance/GDP)
DCPI (inflation)
DGDP (growth rate of real GDP)
MCOV (import cover)
TDPY (total external debt/GDP)
TSPX (total external debt service/exports)
RYPC (real GDP/population

The second and third approaches to variable selection were both based on the results of the principal components analysis: in the one case, the first 'n' component scores were used as inputs to the cluster analysis, while in the other a set of variables was chosen on the basis that each member of the set is

maximally loaded on one of the first 'n' principal components. Nineteen components had eigenvalues exceeding unity, but a set of nineteen scores or variables would not constitute a 'parsimonious' set. The first twelve components, which account for two-thirds of the total variance within the data set, also represent the major economic and financial dimensions of the sample, so further analysis focused on the first twelve components.

The twelve variables that are listed in Table V.3 are maximally loaded on the first twelve components, in the order shown:

Table V.3: Alternative selection of clustering variables: based on PCA

1 TSPY	5 DGDP	9 TSPDG
2 DCPI	6 CARA	10 ILMAG
3 EFMT	7 POPN	11 RTDC
4 AGRP	8 TDRA	12 RYPCG

Only three of these variables (CARA, DCPI and DGDP) appear in the earlier list (Table V.2).

V.3.3 The cluster analysis

The analysis was carried out both for the complete sample of 440 country-year cases, and also on a

year-by-year basis. In all cases, the variables were standardised to zero mean and unit variance.

Clusters were sought on two different criteria, in all cases by examining the dendrograms. First, the year-by-year sub-samples were used to look for clusters of countries whose membership remained fairly stable over time. Secondly, both the full samples and the sub-samples were used to look for clusters of 'arrears' cases or of 'non-arrears' cases. These approaches turned out to be related, because about half the countries in the sample were always in the 'non-arrears' category. In each case a cluster dated t was associated with countries' debt-servicing status at t+1.

Clustering was attempted using all seven variables from Table V.2. Next, the first twelve component scores were used, and then the first seven. Thirdly, the twelve variables from Table V.3 were used: initially all of them, then the first four, first six and first seven were tried. It became clear that Ward's method, applied to the seven variables listed in Table V.2, yielded clusters that were more distinct and that were more stable over time, compared with clusters formed on any other basis.

The dendrograms for this cluster analysis are set out in Figure V.1. In each year a well defined cluster exists in which there is a core membership that remains fairly stable from year to year. This core consists of countries that are diverse in many respects, but a common feature is that over the sample period most of the clustered country-year cases were 'non-arrears'. This cluster is referred to as 'cluster I'.

Figure V.1: Dendrograms from cluster analysis

Variables used (after standardisation to zero mean and unit standard deviation):

CARA current account balance/GDP

DCPI consumer price inflation

DGDP annual growth rate of GDP

MCOV import cover

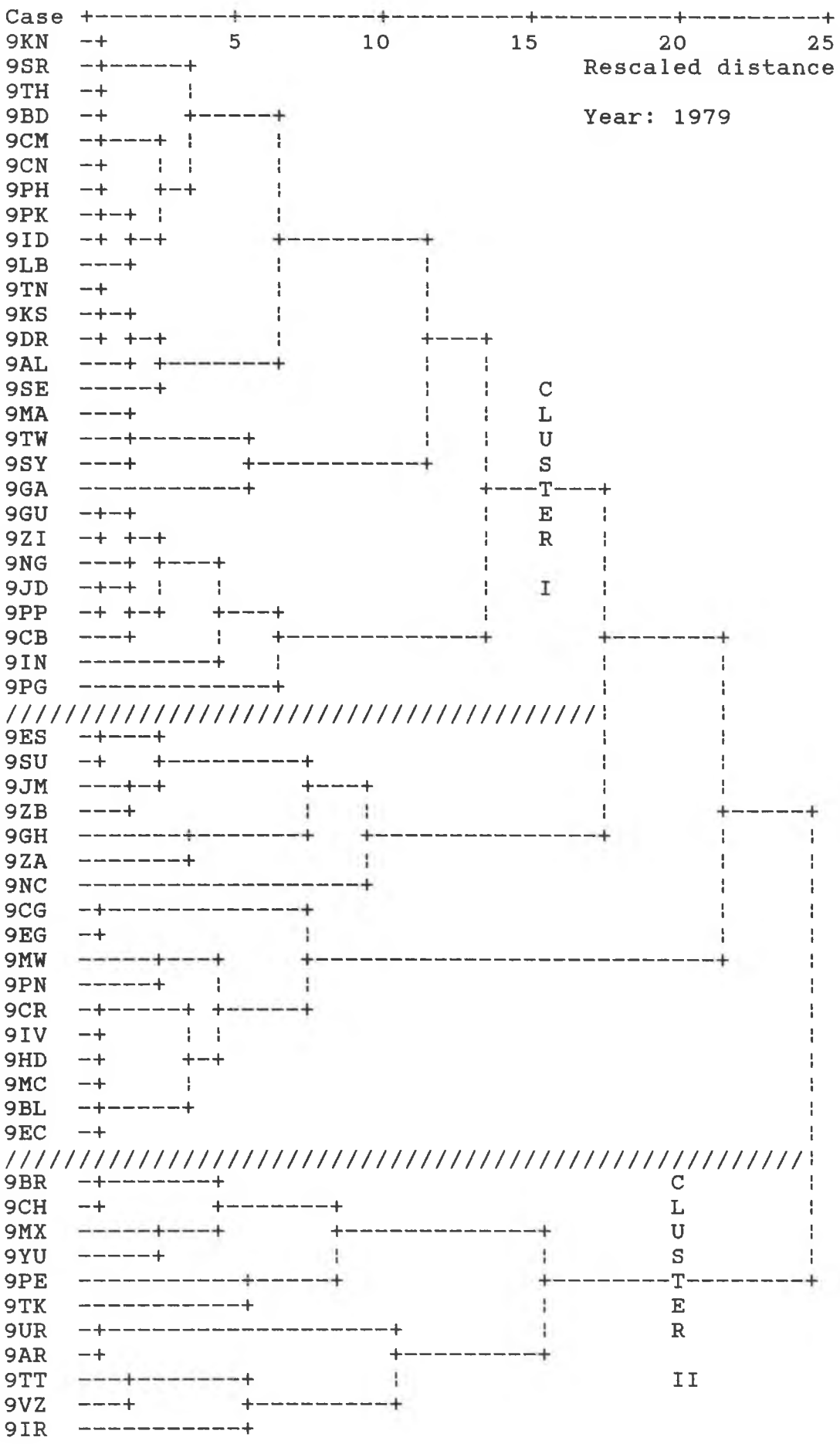
TDPY total external debt/GDP

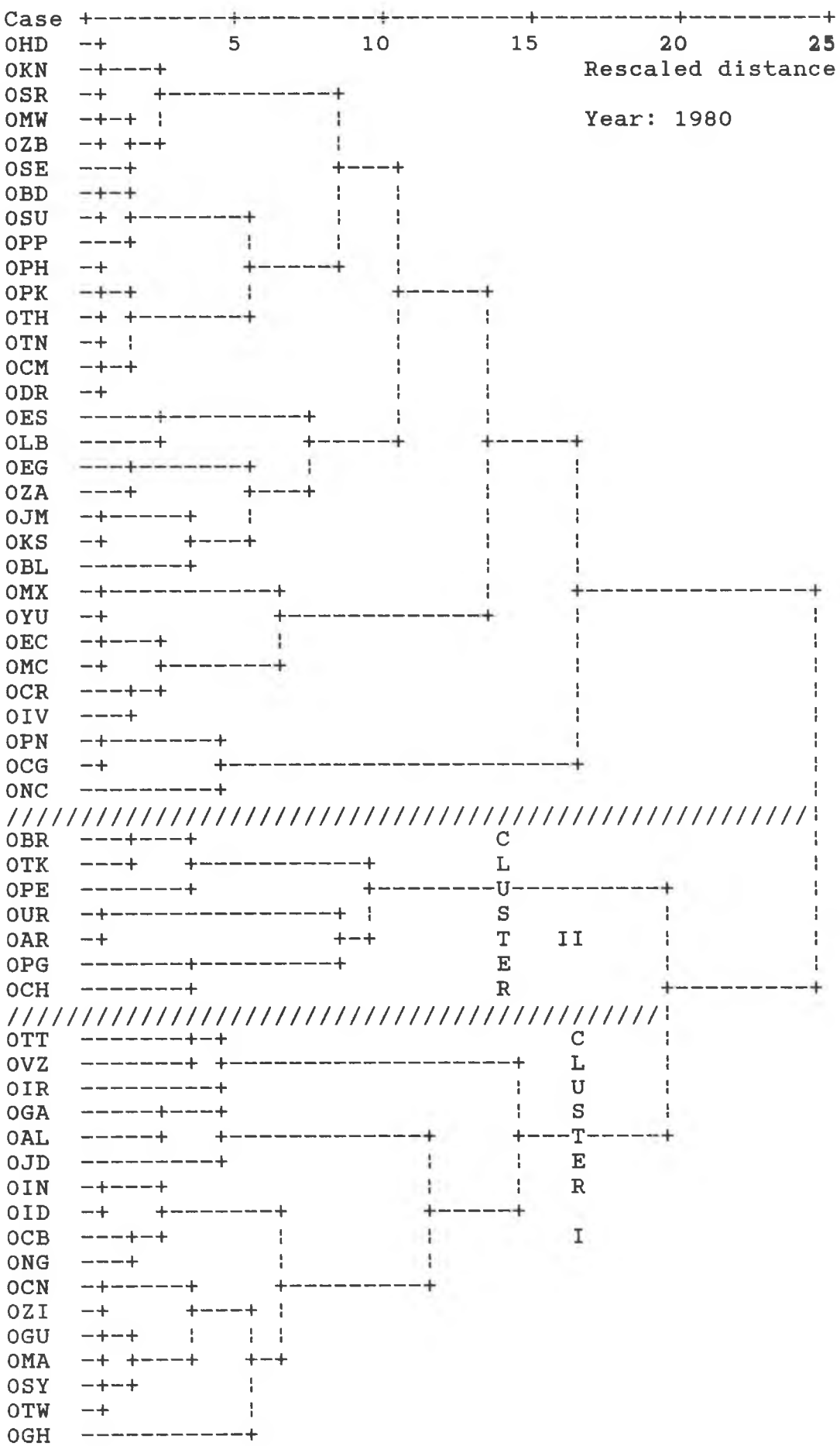
TSPX total external debt service/exports

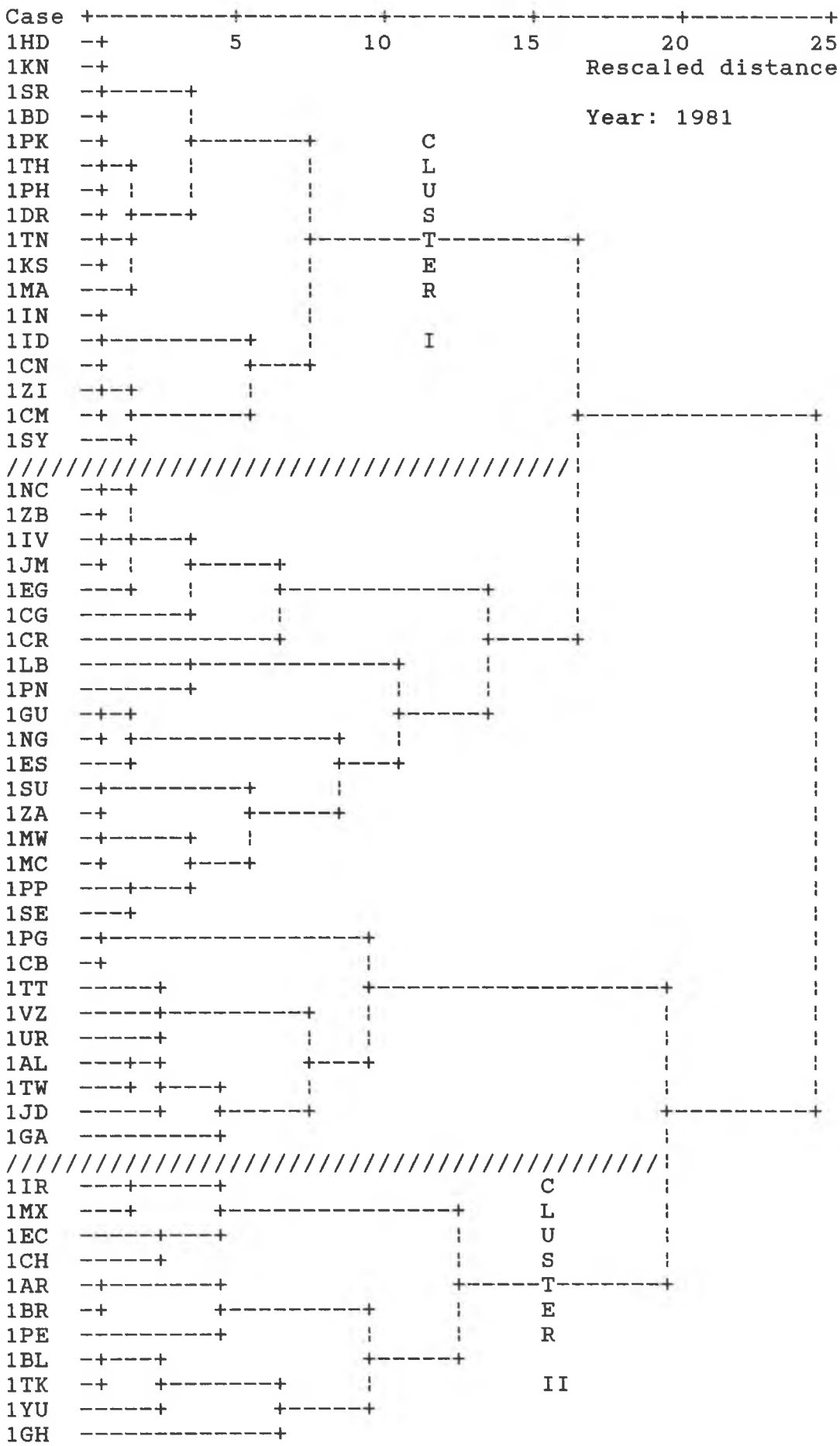
RYPC real income per head of population

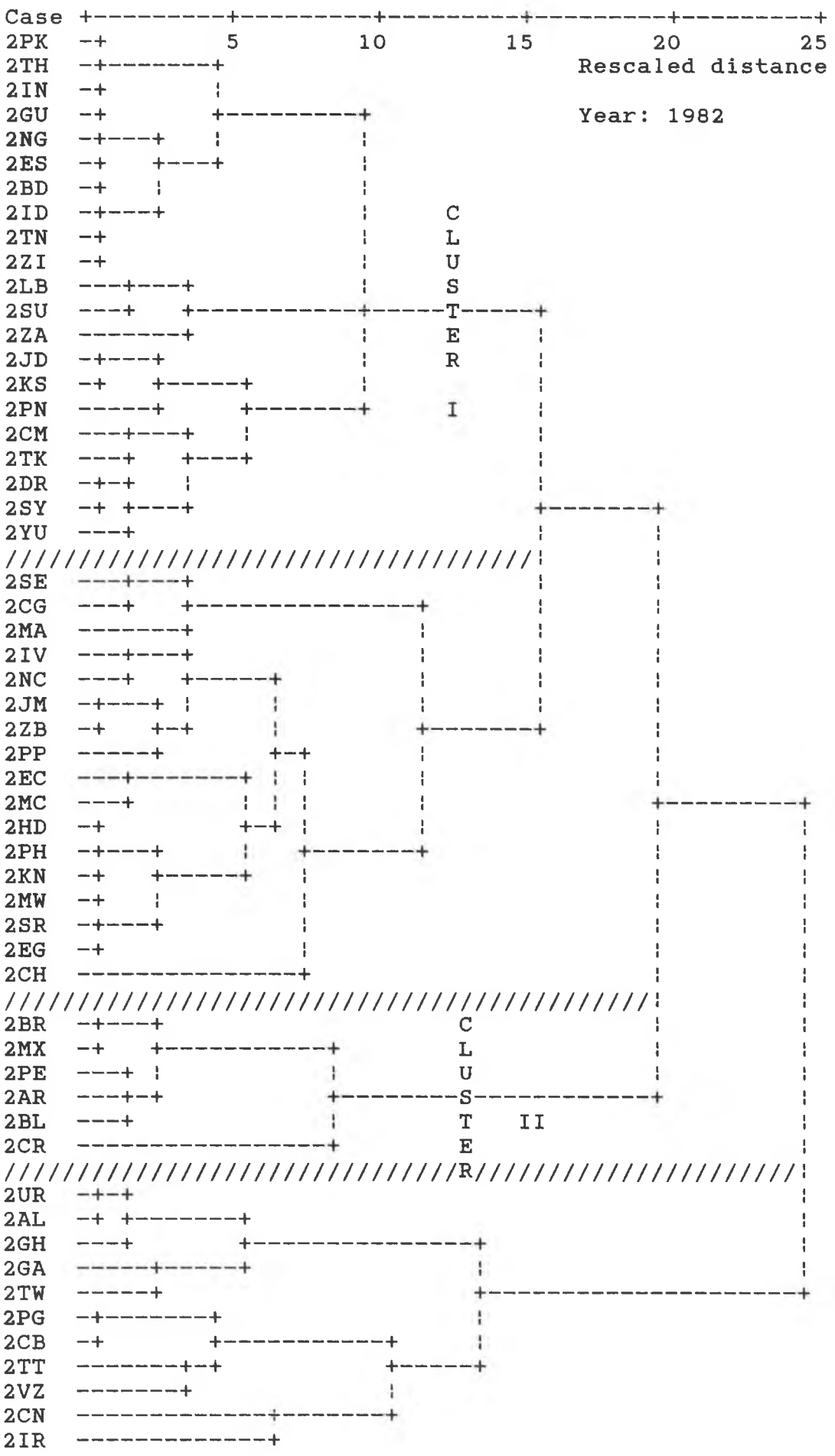
Separate analyses are presented for each sample year, 1979 - 1986, using Ward's method of clustering. The dendrograms for each year follow this page. The codings refer to the year [(197)9 to (198)6], followed by the label identifying the country.

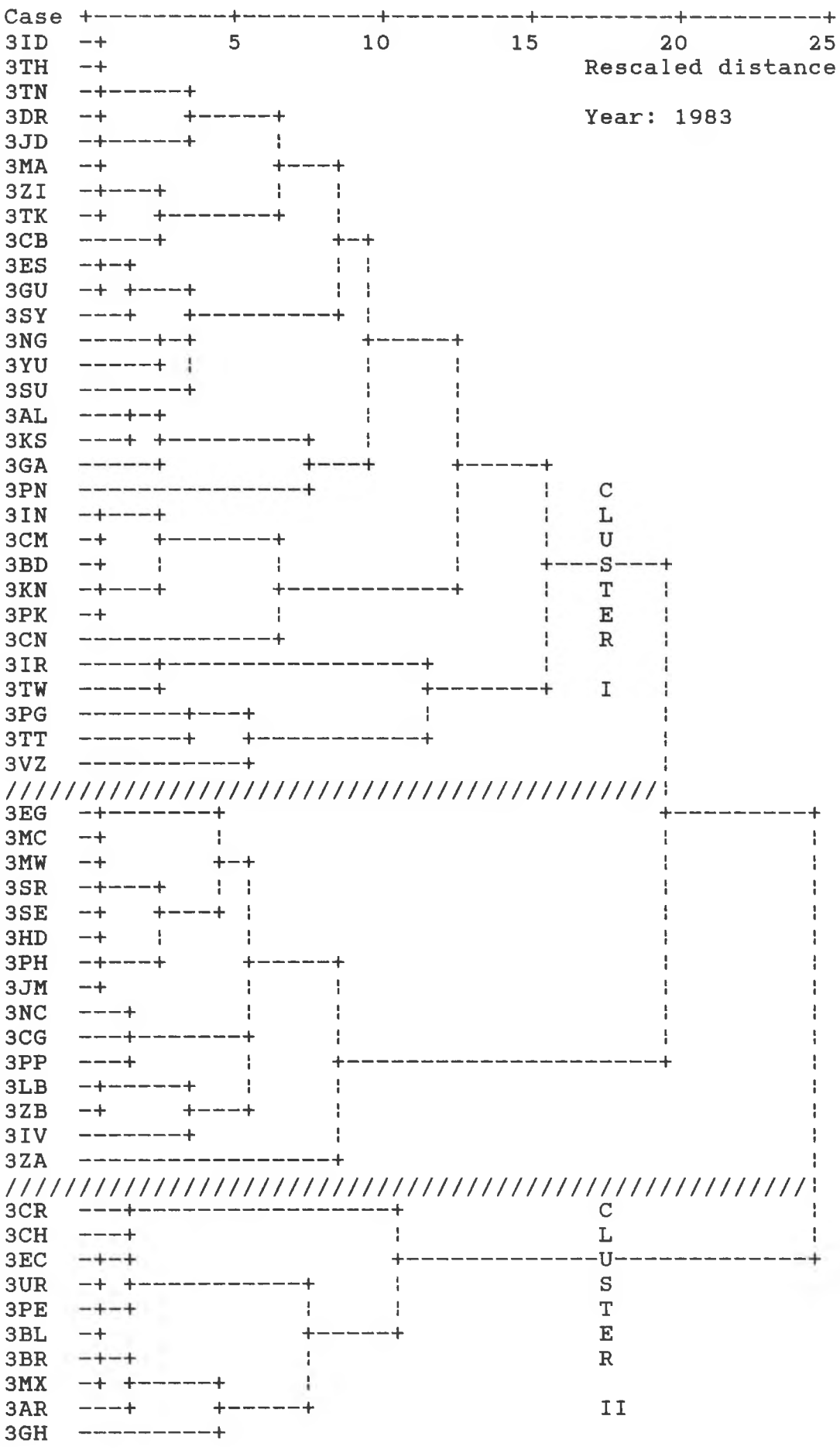
.../

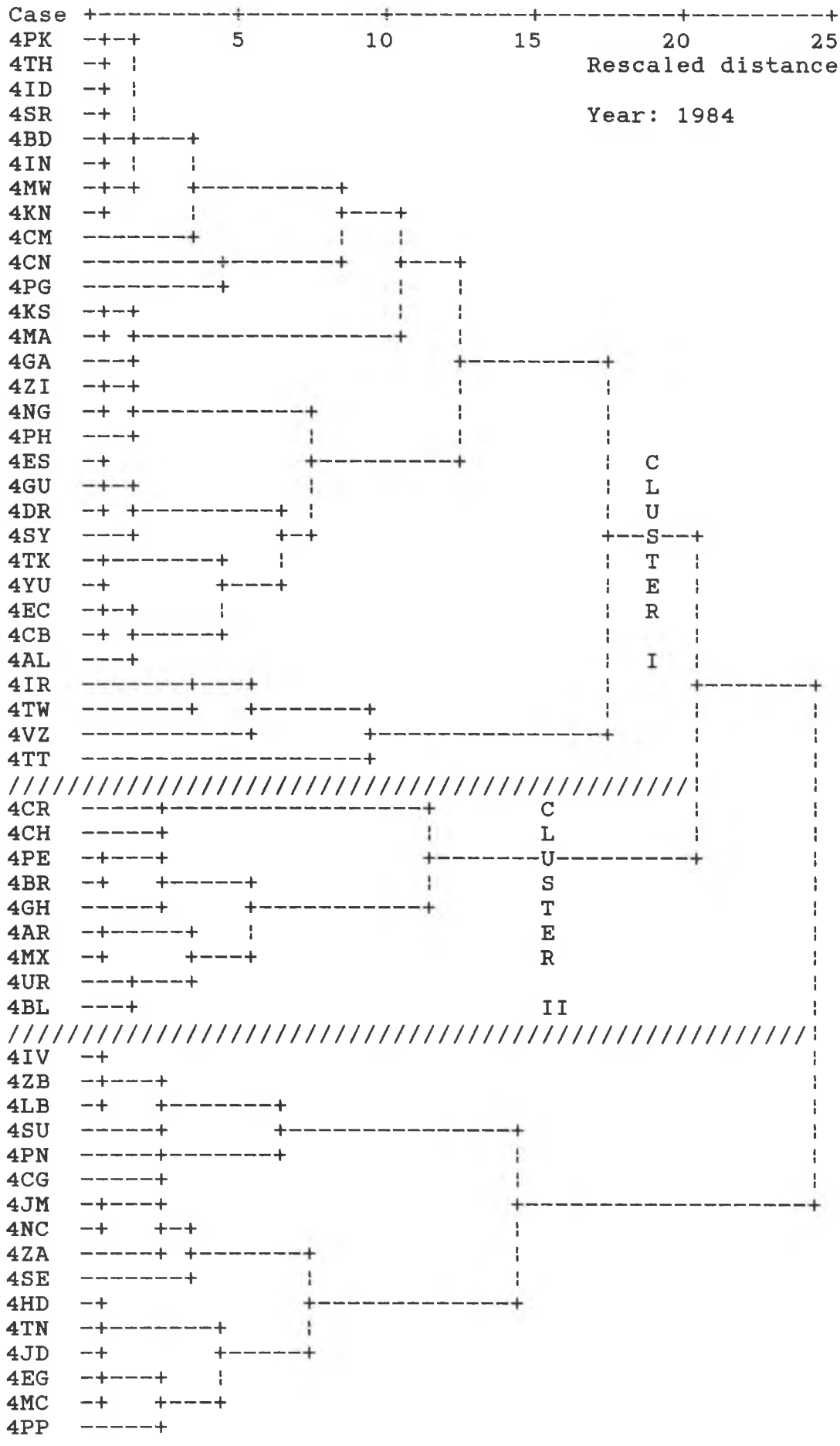

Case
9KN
9SR
9TH
9BD
9CM
9CN
9PH
9PK
9ID
9LB
9TN
9KS
9DR
9AL
9SE
9MA
9TW
9SY
9GA
9GU
9ZI
9NG
9JD
9PP
9CB
9IN
9PG
9ES
9SU
9JM
9ZB
9GH
9ZA
9NC
9CG
9EG
9MW
9PN
9CR
9IV
9HD
9MC
9BL
9EC
9BR
9CH
9MX
9YU
9PE
9TK
9UR
9AR
9TT
9VZ
9IR
5
10
15
20
25
Rescaled distance
Year: 1979
C
L
U
S
T
E
R
I
C
L
U
S
T
E
R
II

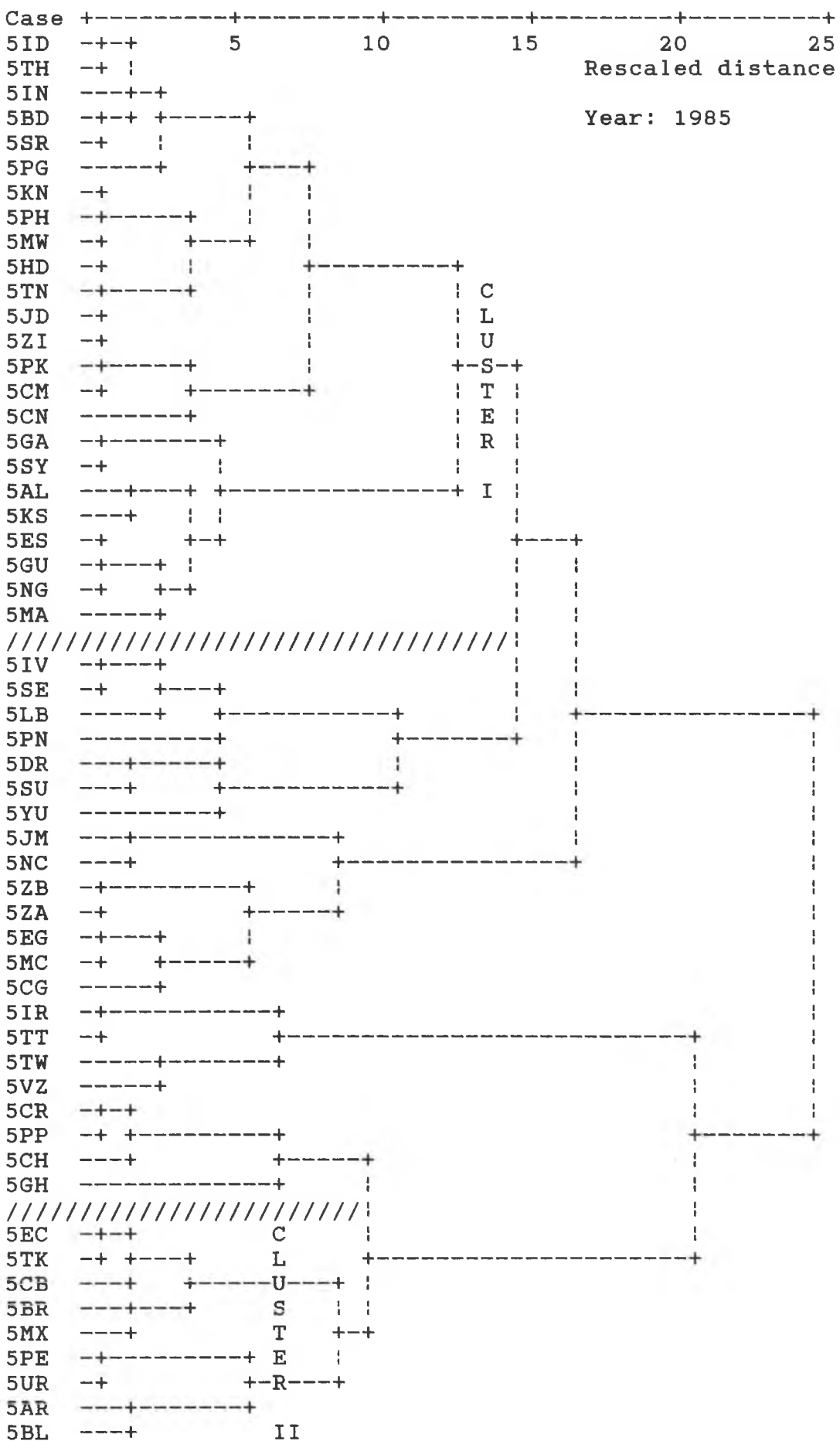

Case
OHD
OKN
OSR
OMW
OZB
OSE
OBD
OSU
OPP
OPH
OPK
OTH
OTN
OCM
ODR
OES
OLB
OEG
OZA
OJM
OKS
OBL
OMX
OYU
OEC
OMC
OCR
OIV
OPN
OCG
ONC
OBR
OTK
OPE
OUR
OAR
OPG
OCH
OTT
OVZ
OIR
OGA
OAL
OJD
OIN
OID
OCB
ONG
OCN
OZI
OGU
OMA
OSY
OTW
OGH
5
10
15
20
25
Rescaled distance
Year: 1980
CLUSTER II

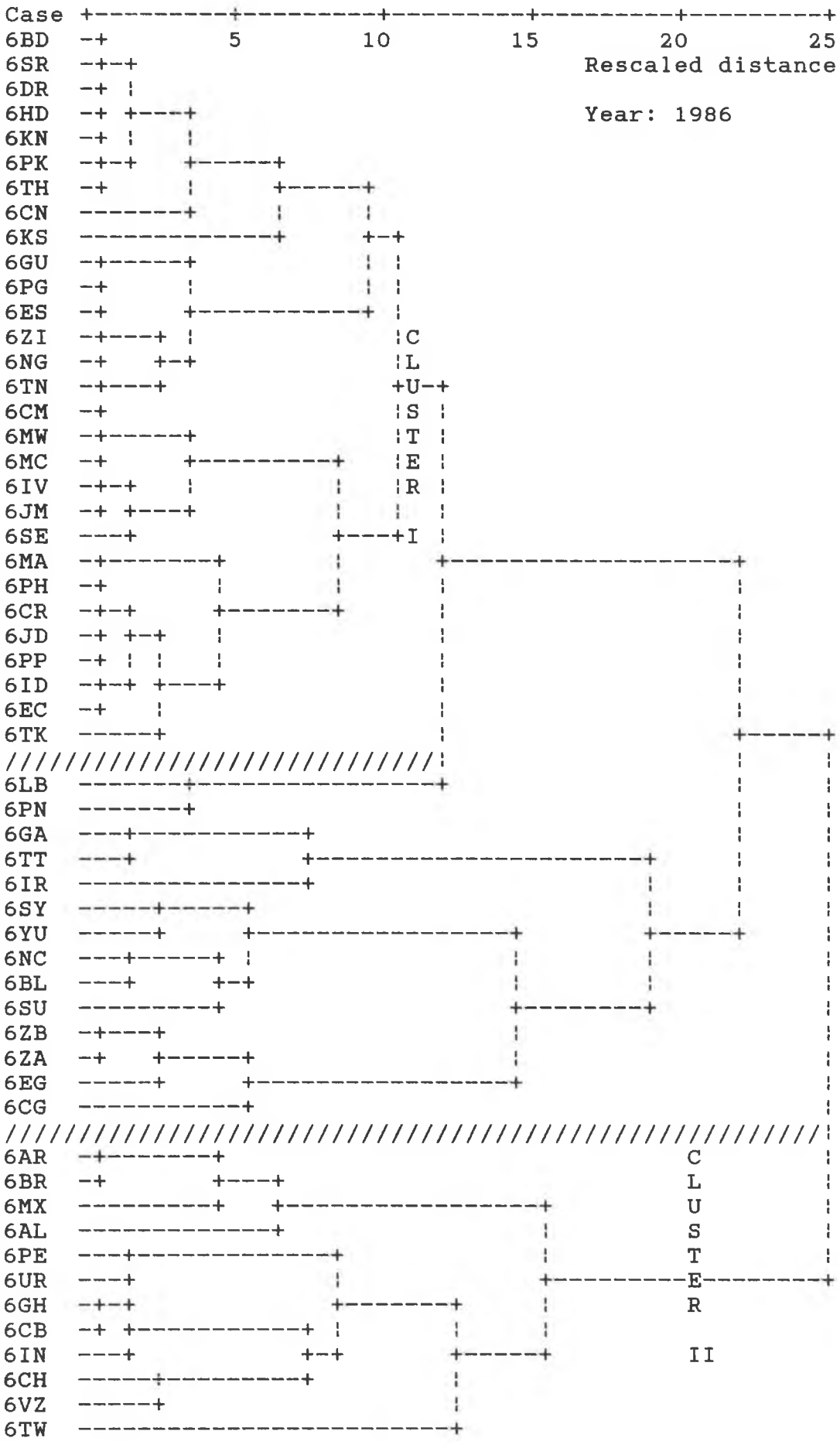

Case
1HD
1KN
1SR
1BD
1PK
1TH
1PH
1DR
1TN
1KS
1MA
1IN
1ID
1CN
1ZI
1CM
1SY
1NC
1ZB
1IV
1JM
1EG
1CG
1CR
1LB
1PN
1GU
1NG
1ES
1SU
1ZA
1MW
1MC
1PP
1SE
1PG
1CB
1TT
1VZ
1UR
1AL
1TW
1JD
1GA
1IR
1MX
1EC
1CH
1AR
1BR
1PE
1BL
1TK
1YU
1GH
5
10
15
20
25
Rescaled distance
Year: 1981
C
L
U
S
T
E
R
I
C
L
U
S
T
E
R
II


Case
2PK
2TH
2IN
2GU
2NG
2ES
2BD
2ID
2TN
2ZI
2LB
2SU
2ZA
2JD
2KS
2PN
2CM
2TK
2DR
2SY
2YU
2SE
2CG
2MA
2IV
2NC
2JM
2ZB
2PP
2EC
2MC
2HD
2PH
2KN
2MW
2SR
2EG
2CH
2BR
2MX
2PE
2AR
2BL
2CR
2UR
2AL
2GH
2GA
2TW
2PG
2CB
2TT
2VZ
2CN
2IR
5
10
15
20
25
Rescaled distance
Year: 1982
C
L
U
S
T
E
R
I
C
L
U
S
T
E
R
II


Case
3ID
3TH
3TN
3DR
3JD
3MA
3ZI
3TK
3CB
3ES
3GU
3SY
3NG
3YU
3SU
3AL
3KS
3GA
3PN
3IN
3CM
3BD
3KN
3PK
3CN
3IR
3TW
3PG
3TT
3VZ
3EG
3MC
3MW
3SR
3SE
3HD
3PH
3JM
3NC
3CG
3PP
3LB
3ZB
3IV
3ZA
3CR
3CH
3EC
3UR
3PE
3BL
3BR
3MX
3AR
3GH
5
10
15
20
25
Rescaled distance
Year: 1983
C
L
U
S
T
E
R
I
C
L
U
S
T
E
R
II


Case
4PK
4TH
4ID
4SR
4BD
4IN
4MW
4KN
4CM
4CN
4PG
4KS
4MA
4GA
4ZI
4NG
4PH
4ES
4GU
4DR
4SY
4TK
4YU
4EC
4CB
4AL
4IR
4TW
4VZ
4TT
4CR
4CH
4PE
4BR
4GH
4AR
4MX
4UR
4BL
4IV
4ZB
4LB
4SU
4PN
4CG
4JM
4NC
4ZA
4SE
4HD
4TN
4JD
4EG
4MC
4PP
5
10
15
20
25
Rescaled distance
Year: 1984
C
L
U
S
T
E
R
I
C
L
U
S
T
E
R
II


Case
5ID
5TH
5IN
5BD
5SR
5PG
5KN
5PH
5MW
5HD
5TN
5JD
5ZI
5PK
5CM
5CN
5GA
5SY
5AL
5KS
5ES
5GU
5NG
5MA
5IV
5SE
5LB
5PN
5DR
5SU
5YU
5JM
5NC
5ZB
5ZA
5EG
5MC
5CG
5IR
5TT
5TW
5VZ
5CR
5PP
5CH
5GH
5EC
5TK
5CB
5BR
5MX
5PE
5UR
5AR
5BL
5
10
15
20
25
Rescaled distance
Year: 1985
C
L
U
S
T
E
R
I
C
L
U
S
T
E
R
II


Case
6BD
6SR
6DR
6HD
6KN
6PK
6TH
6CN
6KS
6GU
6PG
6ES
6ZI
6NG
6TN
6CM
6MW
6MC
6IV
6JM
6SE
6MA
6PH
6CR
6JD
6PP
6ID
6EC
6TK
6LB
6PN
6GA
6TT
6IR
6SY
6YU
6NC
6BL
6SU
6ZB
6ZA
6EG
6CG
6AR
6BR
6MX
6AL
6PE
6UR
6GH
6CB
6IN
6CH
6VZ
6TW
5
10
15
20
25
Rescaled distance
Year: 1986
C
L
U
S
T
E
R
II

V.3.4 Cluster analysis: conclusions

A summary of the analysis is presented in Table V.4. The first eighteen countries in the list were clustered four times or more, from a potential maximum of eight years, and moreover were never 'arrears' or 'weak-year' cases during the sample period, except that India had a weak-year in 1980 (relating to the 1979 cluster).

A further twenty-six countries featured in between one and seven clusters. Their order in the table reflects their standing as debtors: those listed earlier are 'non-arrears' countries, then come countries with some weak years, and finally countries with some 'arrears' years.

The next question concerns the interpretation of the clusters. One approach is to look for common characteristics of the countries thus grouped, independently of the clustering variables. The prevalence of non-arrears cases in cluster I raises the possibility that the cluster analysis might be identifying countries that are consistently in that category during the sample period. This in turn suggests that the analysis might be used to construct an 'early warning' indicator; this matter will be taken up in chapter VI.

Table V.4: Membership of cluster I

Key: Status at t+1, in column for cluster at t:
**: non-arrears; wy: weak; aa: arrears;
blank entry: country not clustered in year

Case	1979	1980	1981	1982	1983	1984	1985	1986
ZI	**	**	**	**	**	**	**	**
ID	**	**	**	**	**	**	**	**
BD	**		**	**	**	**	**	**
CM	**		**	**	**	**	**	**
KS	**		**	**	**	**	**	**
TH	**		**	**	**	**	**	**
CN	**	**	**		**	**	**	**
MA	**	**	**		**	**	**	**
SY	**	**	**	**	**	**	**	
KN	**		**		**	**	**	**
TN	**		**	**	**		**	**
JD	**	**		**	**		**	**
IN	wy	**	**	**	**	**	**	
AL	**	**			**	**	**	
ES				**	**	**	**	**
SR	**		**			**	**	**
CB	**	**			**	**		
TW	**	**			**	**		
IR		**			**	**		
TT		**			**	wy		
PP	**							**
PG	**				wy	wy	wy	aa
GU	**	**		wy	wy	wy	aa	wy
GH		wy						
MW						wy	wy	aa
TK				wy	wy	wy		**
PK	wy		wy	wy	wy	**	**	**
CR								wy
IV								aa
SE	wy							aa
VZ		wy			wy	wy		
HD			wy				wy	wy
MC								aa
JM								aa
ZA				aa				
EC						aa		aa
PN				aa	wy			
GA	wy	wy			wy	wy	aa	
PH	wy		wy			wy	wy	aa
SU				aa	aa			
LB	aa			aa				
NG	wy	wy		aa	wy	wy	aa	wy
YU				aa	aa	aa		
DR	wy		aa	wy	aa	aa		wy

Not clustered: AR BL BR CG CH EG MX NC PE UR ZB

An alternative approach is to focus on the clustering variables, information about whose sample means is set out in Table V.5.

Table V.5: Sample means for cluster I

		Variables						
Year	Countries	CARA	DCPI	DGDP	MCOV	TDPY	TSPX	RYPC
All	All	-4.1	20.6	2.8	2.6	54.6	23.8	1199.9
1979	Cluster	-1.9	11.6	5.5	3.3	31.0	13.8	1035.8
1979	Non-cluster	-3.9	28.0	2.9	2.8	48.7	25.6	1570.6
1979	All	-2.9	19.9	4.2	3.0	40.0	19.8	1308.0
1980	Cluster	1.8	16.3	6.1	3.8	22.7	11.3	1656.6
1980	Non-cluster	-7.5	25.7	3.6	2.0	49.9	23.6	1174.0
1980	All	-4.7	22.8	4.4	2.6	41.5	19.8	1323.2
1981	Cluster	-5.5	12.4	6.9	1.9	35.0	15.3	762.0
1981	Non-cluster	-6.7	23.5	1.8	2.2	50.5	26.0	1532.0
1981	All	-6.3	20.1	3.4	2.1	45.7	22.7	1294.0
1982	Cluster	-3.7	12.8	1.9	1.3	43.9	16.8	910.8
1982	Non-cluster	-7.1	21.4	1.0	2.7	56.5	29.1	1398.0
1982	All	-5.8	18.1	1.3	2.1	51.7	24.4	1212.0
1983	Cluster	-2.4	12.3	2.9	2.9	38.9	19.8	1328.8
1983	Non-cluster	-7.5	30.9	-0.3	2.2	82.0	32.4	973.1
1983	All	-4.7	20.7	1.5	2.5	58.5	25.5	1167.1
1984	Cluster	-1.1	16.2	3.0	2.9	38.1	21.0	1215.0
1984	Non-cluster	-5.2	28.4	1.8	2.3	90.3	27.5	939.9
1984	All	-3.0	21.7	2.5	2.6	61.8	24.0	1090.0
1985	Cluster	-3.0	9.3	4.0	2.1	47.6	24.0	896.2
1985	Non-cluster	-2.5	29.1	1.2	3.2	84.9	28.2	1182.2
1985	All	-2.7	20.5	2.4	2.7	68.6	26.4	1057.4
1986	Cluster	-1.9	10.8	3.4	2.2	66.5	27.9	787.8
1986	Non-cluster	-4.0	32.4	1.3	3.5	71.0	27.3	1548.4
1986	All	-2.9	21.0	2.4	2.8	68.6	27.6	1147.3

If the sample means for the clusters are compared with those for the non-cluster countries, or for the

whole sample, a ready interpretation of the cluster emerges. The clustered countries have on average:

higher: current account/GDP (except 1985)
economic growth

lower: inflation
total debt/GDP
lower total debt service/exports (except 1986)

The comparisons of import cover and real income/head show no such consistent patterns. For this reason, cluster analyses were run in which each and both were excluded, but no interpretable clusters emerged.

In summary, Cluster I consists of a group of countries that have not, on the whole, had debt-servicing difficulties. That feature of the cluster also implies, of course, that it does not yield a partition of the sample that could be used in estimating early-warning models, because each group formed by such a partition must include a reasonable representation of countries in arrears.

A second moderately stable cluster appears in each year, and is referred to as 'cluster II'. It has a 'core' membership of Argentina, Brazil, Mexico and Peru. Except for Mexico in 1980, these countries always appear. Other sometime members are Chile and Uruguay (six appearances each), Bolivia (five) Turkey (four), Costa Rica, Ecuador and Ghana (three each),

and Venezuela and Yugoslavia (twice each). Some of the largest problem debtors are captured by cluster II. However, there is no consistent clustering of problem countries in their 'arrears' years. Conversely, in some years the cluster picks up countries which have not been in difficulties, such as Algeria, Colombia and Taiwan. The factor that unifies most of the members of this cluster in most years is geographical - it is largely a Latin American cluster. This is a factor that will be handled in later analysis by way of a regional dummy variable.

VI Partitions within the data set: applications to 'early warning'

VI.1 Introduction

A primary purpose of this thesis is to derive operational models for providing early warning of difficulties with debt-servicing. Various multivariate statistical techniques may be appropriate for this, and they are explored in this chapter and chapters VII to X. The results are evaluated in chapter XI, in terms of the extent to which the various models provide true *ex ante* prediction.

An alternative approach is to rely on expert judgement, and after exploring this possibility in chapters XII and XIII, the results of the two approaches are compared in chapter XIV. Finally, the research findings are summarised in chapter XV.

The multivariate statistical approach to early warning is introduced in this chapter. First, the cluster analysis of chapter V is examined for evidence of 'early warning' capability. Secondly, a proximities analysis is derived from the cluster analysis, and examined in a similar fashion.

Schmidt (1984) has applied clustering methods to the early-warning problem. The logic of this approach is that arrears countries may be close to each other in the multidimensional space of key financial and economic variables; by the same token, non-arrears countries may also be close to each other. The cluster analysis attempts to find distinct clusters of country-year cases, in which the arrears and non-arrears cases are separated.

VI.2 Cluster analysis

In chapter V, a method of clustering was found that yielded a distinct cluster, that was present in each sample year, and that had a fairly constant membership of non-arrears countries. In this section, the cluster (cluster I) is further examined as an early-warning indicator. A summary of the cluster analysis is presented in Table V.4. The first eighteen countries in the list are clustered four times or more, from a potential maximum of eight years. These countries were never 'arrears' or 'weak-year' cases during the sample period, other than India which had a weak-year in 1980.

As reported in V.3.4, a further twenty-six countries feature in between one and seven clusters, with their order in Table V.4 reflecting their standing as debtors. Of twenty-four countries that cluster when 'weak' or 'arrears', seventeen do so no more than three times in that condition. The other seven countries, with the condition in which they appear, are set out in Table VI.1.

Table VI.1: Cluster I: summary of repeated misclassifications

Pakistan	: 4 weak years (also 3 non-arrears)
Paraguay	: 3 weak years, 1 arrears (also 1 non-arrears)
Guatemala	: 4 weak-years, 1 arrears (also 2 non-arrears)
Gabon) each Philippines)	: 4 weak years, 1 arrears
Dominican Republic	: 3 weak years, 3 arrears
Nigeria	: 5 weak years, 2 arrears

Note:
Only countries with at least four misclassifications are shown in the table.

If we interpret the cluster as essentially consisting of non-arrears cases, then these seven are the only countries to be consistently misclassified. Misclassification of arrears years is the more costly variety, and the incidence of this is fairly high, particularly in 1983 and 1987, as Table VI.2 reveals.

In each year, the group of countries that is excluded from cluster I does not itself form an identifiable and stable cluster, so that the analysis does not partition the sample into 'good' and 'bad' credit risks. The high type II error rate illustrates this point. It equals 48 per cent or more in 1981, 1982

and 1983, and it lies below 20 per cent only in 1984 (10 per cent) and 1985 (14 per cent). The only published work that reports the application of clustering methods to country risk analysis is Schmidt (1984), who reports a type II error rate of 73 per cent in one of his best examples. While the error rates reported here are lower than that, they are still unacceptably high.

Table VI.2: Summary of cluster I membership

Cluster date:	1979	1980	1981	1982	1983	1984	1985	1986
Total members	27	17	17	21	30	30	24	29
Of which debt-servicing status in the following year:								
non arrears	19	13	13	11	19	18	17	16
weak	7	4	3	4	8	9	4	5
arrears	1	0	1	6	3	3	3	8
T1 error rate	25%	0%	9%	26%	23%	18%	25%	53%
Cluster date:	**1979**	**1980**	**1981**	**1982**	**1983**	**1984**	**1985**	**1986**
Non-members	28	38	38	34	25	25	31	26
Of which debt-servicing status in the following year:								
non arrears	5	12	13	11	2	3	5	6
T2 error rate	21%	48%	50%	50%	10%	14%	23%	27%
weak	20	16	15	6	13	8	17	13
arrears	3	10	10	17	10	14	9	7

Note:
Weak-year cases excluded from calculation of error rates.

VI.3 Proximities analysis

The countries clustered together are close in terms of the distance between their points of location in seven-dimensional space, where the dimensions represent the seven variables of Table V.2, after normalisation to zero mean and unit standard deviation. Unlike discriminant and logit analyses, clustering does not generate a 'function' that can be applied to the problem of out-of-sample classification. However, clustering is based on the proximity of cases in multi-dimensional space, and a proximities analysis can be made the basis for out-of-sample classification.

To prepare the ground for the out-of-sample analysis of chapter XI, a proximities analysis is undertaken here, separately for each year. First, the sample centroid for the clustered countries is computed. Then the Euclidean distance from the centroid is computed for each member of the sample. This is a limited analysis: it focuses on the distances between each case and the cluster centroid, and ignores the distances between cases. This is because the purpose is to develop and evaluate a reference point (the cluster centroid) for out-of-sample testing.

Three possibilities arise for computing cluster means. The most obvious is to compute them for each year over all countries clustered in that year. The second excludes weak-year and arrears countries in years when they fall in either category. The third stabilises the second by using a constant group of non-arrears countries that are clustered in most years. 'Most' is taken to mean at least four, and leads to the choice of the first eighteen cases listed in Table V.4.

Since the purpose of the analysis is to construct a reference point for basing a judgement about a country's credit-worthiness, it is advantageous to derive it from a constant grouping of countries, and moreover from a grouping that includes only non-arrears cases. When the joint means of this constant-membership cluster are compared with those of the rest of the sample, it turns out that average real income/head is lower within the cluster, while import cover is higher. Along the other five dimensions of clustering, the relative positions of the cluster means are the same as already described for the full clusters - i.e. higher current account balance/GDP, lower inflation, higher GDP growth, and lower debt and debt-service ratios.

The hypothesis proposed is that proximity to the

centroid of cluster I in any year is a sufficient, but not necessary, indicator of absence of debt-servicing difficulties, in the sense that the country will in fact be 'non arrears' in the following year. A cutoff of two standard deviations is used to determine whether cases are to be classified as close to or far from the centroid. This distance is arbitrary, but inspection of the annual sets of Euclidean distances revealed that it yields a reasonable cut between non-arrears and other cases, one year ahead. On this basis, the number of weak-year or arrears cases classified as 'close' to the centroid of the cluster ranges from three cases in 1981 (using 1980 proximities) to eight in 1987 (using 1986 proximities). However, most of the misclassifications are of weak-year cases. Three arrears cases were close to the centroid in 1987, and two in 1983 and 1986 (based on the proximity measure dated one year earlier). Otherwise there was only one such case in each year. These results are summarised in Table VI.3.

Proximity to the cluster centroid is not necessary for a country to be a good credit risk: in other words, it is to be expected that some countries will be remote and yet be credit worthy. This is essentially because the non-cluster group does not itself form an identifiable and stable cluster, so

Table VI.3: Results of proximities analysis

Proximity dated:	1979	1980	1981	1982	1983	1984	1985	1986
Countries within two standard deviations of centroid of constant-membership cluster I:								
Numbers	20	17	19	19	18	17	22	23
of which debt-servicing status in following year:								
weak-year cases	3	2	3	4	4	3	5	5
arrears	1	1	1	2	1	1	2	3
T1 error rates	25%	10%	9%	8%	8%	6%	17%	20%
Proximity dated:	**1979**	**1980**	**1981**	**1982**	**1983**	**1984**	**1985**	**1986**
Countries beyond two standard deviations from centroid:								
Numbers	35	38	36	36	37	38	33	32
of which debt-servicing status in following year:								
non arrears:	8	11	11	9	8	8	7	7
T2 error rate:	33%	44%	30%	29%	28%	28%	24%	24%
Memo: debt-servicing status in following years: totals:								
Weak:	27	20	18	9	22	17	21	18
Arrears:	4	10	11	24	12	17	12	15
Non-arrears:	24	25	26	22	21	21	22	22

Notes:
Weak-year cases excluded from calculation of error rates.

that the analysis does not partition the sample into 'good' and 'bad' credit risks. Not only do certain non-arrears countries consistently fall into the lower part of the ranking: some of them, including countries that define the cluster, such as Taiwan, frequently lie toward the outer limit of distance from the centroid. The point here is that cases that are distant from the cluster centroid may, at the

same time, be remote from each other. Type II error rates range between 24 per cent and 44 per cent. They differ from those reported for the cluster analysis in part because of the arbitrary cutoff of 2 standard deviations that was used for classification purposes.

In conclusion, the best that may be claimed for the cluster analysis is that it can identify a set of countries that are broadly non-arrears. The incidence of Type II errors is high, although lower than that reported in a similar study by Schmidt (1984). The proximities analysis, whose purpose is to make the cluster analysis operational for out-of-sample classification, yields a similarly high incidence of Type II errors. These results are not promising, and provide a motive for turning to other methods of classification. However, a final assessment of clustering and proximities analyses awaits the analysis of out-of-sample performance in chapter XI.

Appendix VI

Table VIA.1: Sample means for constant-membership cluster of eighteen countries

Year	Countries	Variables CARA	DCPI	DGDP	MCOV	TDPY	TSPX	RYPC
All	Cluster	-2.9	11.4	4.4	2.8	35.5	17.7	1035.6
All	Other	-4.8	25.9	1.8	2.4	65.7	27.2	1293.7
All	All	-4.1	20.6	2.8	2.6	54.6	23.8	1199.9
1979	Cluster	-1.7	12.3	5.5	3.5	27.6	12.2	1022.1
1979	Other	-3.6	24.3	3.4	2.8	47.1	24.1	1471.5
1979	All	-2.9	19.9	4.2	3.0	40.0	19.8	1308.0
1980	Cluster	-3.0	15.3	6.2	2.9	29.3	12.2	1051.1
1980	Other	-5.6	27.1	3.3	2.4	48.5	24.1	1478.7
1980	All	-4.7	22.8	4.4	2.6	41.5	19.8	1323.2
1981	Cluster	-4.7	13.7	5.9	2.5	31.1	13.9	1032.7
1981	Other	-7.3	23.7	1.9	1.8	54.1	27.7	1443.4
1981	All	-6.3	20.1	3.4	2.1	45.7	22.7	1294.0
1982	Cluster	-4.2	9.7	2.0	2.7	34.9	16.6	1016.8
1982	Other	-6.8	22.9	0.9	1.8	61.3	28.9	1323.5
1982	All	-5.8	18.1	1.3	2.1	51.7	24.4	1212.0
1983	Cluster	-3.6	9.6	4.2	2.8	37.5	19.4	1030.7
1983	Other	-5.3	27.1	-0.1	2.2	70.5	29.1	1245.0
1983	All	-4.7	20.7	1.5	2.5	58.5	25.5	1167.1
1984	Cluster	-2.2	8.8	4.3	2.7	37.9	19.0	1019.4
1984	Other	-3.4	29.1	1.4	2.5	75.5	26.8	1130.3
1984	All	-3.0	21.7	2.5	2.6	61.8	24.0	1090.0
1985	Cluster	-2.4	9.5	4.2	2.6	41.6	22.8	1020.3
1985	Other	-2.8	26.7	1.4	2.8	84.1	28.4	1078.6
1985	All	-2.7	20.5	2.4	2.7	68.6	26.4	1057.4
1986	Cluster	-1.6	12.0	3.1	3.0	44.3	25.7	1092.0
1986	Other	-3.6	26.1	2.0	2.7	82.5	28.7	1178.9
1986	All	-2.9	21.0	2.4	2.8	68.6	27.6	1147.3

VII Early warning models: discriminant analysis

VII.1 Introduction

One of the issues addressed by this thesis is the possibility of forecasting the debt-servicing status of less developed countries. The problem is defined as one of binary classification: one group consists of countries that are servicing their debts in full, while countries in the other group are at the point of emergence of arrears. The objective is to forecast debt-servicing status at date t+1 on the basis of financial and economic characteristics at date t.

The statistical analyses utilise the categorical criterion variable ARS, which is described in section III.8. Several statistical techniques may be applied to this type of problem: one of these is the two-group discriminant analysis of this chapter, while two others are the logit and AID techniques of chapters VIII and IX.

Section VII.2 raises some theoretical issues concerning the use of linear discriminant analysis. In section VII.3, a discriminant model is developed, using the basic data set of 70 continuous financial

and economic variables. Its statistical properties are assessed, as is its within-sample classificatory performance. Section VII.4 reports the results of a discriminant analysis using the basic data set augmented by the categorical variables of Table IV.7. The data set excludes certain cases defined as 'weak' (see chapter III), and the effects of this exclusion are assessed in section VII.5. Finally, in section VII.6, the chapter is summarised and conclusions are drawn

The discriminant analyses utilise the SPSS-X programme DISCRIMINANT and the BMDP programme 7M. The Lachenbruch holdout test and the estimation of jackknifed coefficients utilise programmes that were written by the author using FORTRAN-77.

VII.2 Linear discriminant analysis: theory

Consider a general problem of classification into two groups, indexed 1 and 2. Let x be a vector of predictor (discriminating) variables. Hand (1981: p.71) defines a discriminant function for the two-group case as a general function h(x) leading to the classification rule for any case:

- if $h(x) > k$, then classify the case in group 1
- if $h(x) < k$, then classify the case in group 2

where k is some scalar, appropriately chosen. Moreover, Hand shows that if the conditional probability density functions for x, given membership of groups 1 or 2, are respectively $f(x|1)$ and $f(x|2)$, then h(x) of the form

$$h(x) = f(x|1)/f(x|2)$$

is optimal, using minimum error and minimum risk criteria. In the case where the within-group distributions of x are multivariate normal with equal dispersion matrices, this yields a discriminant function that is linear in x. This is Fisher's linear discriminant function. If the distributions are multivariate normal but the dispersion matrices are unequal, then the optimal form of h(x) is quadratic in x. Alternatively, if the dispersion matrices are equal, Fisher's linear discriminant function maximises the ratio of between-groups

variation to within-groups variation, compared with any other linear function of the predictors, but this function is optimal only under multivariate normality of the predictors within each group.

In conclusion, the optimal decision rule is linear only under the combined assumptions of multivariate normality within each group, and equality of the dispersion matrices. Moreover, these conditions form the maintained hypotheses upon which the standard statistical tests of significance are constructed (Maddala, 1983: p.18). In this chapter, these conditions will be assumed to hold, and the strategy will be to search for a linear discriminant function that will discriminate between arrears and non-arrears cases. The predictor variables will be selected from the data set (see chapter IV), using an interactive stepwise procedure.

In the following sections of this chapter, tests for equality of the dispersion matrices are reported, and the results of these are discussed. As regards normality, it is clear from chapter IV that the complete data set is not multivariate normal. However, what are relevant here are the distributional properties of the variables that enter the discriminant function, and this matter will be investigated below.

VII.3 The basic model: continuous discriminating variables

VII.3.1 The data

Over the sample period, 102 country-year cases have been identified as 'arrears', and 200 as 'non arrears'. The first group consists of country-year cases for which ARS(t+1) = 0, while for the second group ARS(t+1) = 9. The remaining 138 have been identified as 'weak' and are therefore excluded from the sample.

VII.3.2 The cutoff discriminant score

A cutoff value for the discriminant score has to be determined. If the prior probabilities of a case's falling into either group were equal, and if the costs of misclassification were equal, then the cutoff value would be the mid-point between the group centroids. The intercept of the discriminant function (which is arbitrary) is usually set so that this cutoff is zero. Let

$$z = f + dx$$

be the discriminant function, where d and x are vectors of coefficients and discriminating variables.

Let c0 and c1 be the group centroids. If f is set equal to

$$-0.5d(c0 + c1)$$

then clearly any case lying at the mid-point between the group centroids will have a discriminant score of zero.

If the prior probability that a case belongs to the arrears group is P, and if C is the cost of misclassifying an arrears case relative to the cost of misclassifying a non-arrears case, then we define the likelihood ratio:

$$L=PC/(1-P).$$

When the model is set up so that the arrears group is the left-hand group, then we adjust for prior probabilities and misclassification costs by re-defining the intercept as:

$$-0.5d(c0+c1) - \log(L).$$

If $L > 1$, $\log(L) > 0$. This reduces the intercept, so if the zero cutoff is retained, it becomes more likely that a given case will be classified as 'arrears' - as required (Green, 1978: p.174).

The calculation of prior probabilities is based on sample proportions: of 302 cases, the proportions in each group are 0.338 and 0.662. The calculation of C is less straightforward. The penalty arising from a type I error (misclassifying a country that in fact

goes into arrears) arises from unpaid debt service, reflected in a heavy discount in the secondary market. Conversely, when a lender makes a type II error (misclassifying a country that in fact services its debt), a profitable opportunity is missed - as reflected in the interest premium paid by a creditworthy LDC, compared with the rates paid by other safer borrowers. This cost of a type II error will be small relative to the cost of a type I, and C computed as the ratio of these costs would be very large. However, the interest premiums are not a full indication of the perceived risk of lending to LDCs: even the debts of countries such as Colombia and Turkey, that are not in arrears of debt service, trade at a discount in the secondary markets. As a first approximation, secondary market discounts in 1987 are used to compute misclassification costs. The year was chosen on the basis partly of data availability: coverage of all the countries listed below was only available for that year. However, it may also be justified as lying immediately after the sample period. In early 1987, the weighted average discount on the debt of a group of major problem countries, comprising Argentina, Brazil, Chile, Morocco, Mexico, Philippines, Uruguay, Ivory Coast, Venezuela and Yugoslavia was about 35 per cent, while that for two large countries, Algeria and Colombia, that have never been in arrears, was 9.35 per cent.

(World Bank, 1988-1989; Saloman Brothers Inc., 1987. The World Bank's explanation for the existence of discounts on the debt of problem-free countries is 'risk contamination'.) This yields a value for C for the period immediately after the sample period of 35/9.35 or 3.75. This compares with the value of 3 assumed without discussion in Sargen (1977), in the context of a discriminant model of sovereign debt rescheduling. On this basis,

$$P = 102/302 = 0.338$$
$$L = 3.75 \times 0.338/0.662 = 1.91.$$

VII.3.3 Discriminant model I

According to Hendry and Mizon (1985: p.9):

> We seek to design congruent models with parsiminous, near-orthogonal parameterisations precisely because in practice we have found such models to be robust.

A parsimonious set of discriminating variables was sought that would provide a linear discriminant function yielding maximal separation between the two groups. The criterion used was minimization of Wilks's lambda. An interactive stepwise procedure was used to determine the best set of explanatory variables, the purpose of the interaction being the achievement of near-orthogonality. At each stage of the procedure, the 'entering' variable was checked for its level of correlation with each variable

entered previously - if the squared correlation coefficient exceeded 0.5, or if both variables were highly loaded on the same principal component of the data set, then the 'entering' variable was deleted from the data set and the procedure was re-started.

The stepwise procedure was applied to a data set that comprised the 70 primary variables and the 7 environmental variables (see chapter IV). Three primary variables suffer from missing observations. Preliminary runs selected none of these at a reasonably early stage, and therefore they were eliminated from the data set to maximise its size at 302 cases. It was also clear that the environmental variables would not be selected either, but they were retained for later use in defining partitions.

Table VII.1: Discriminant model I

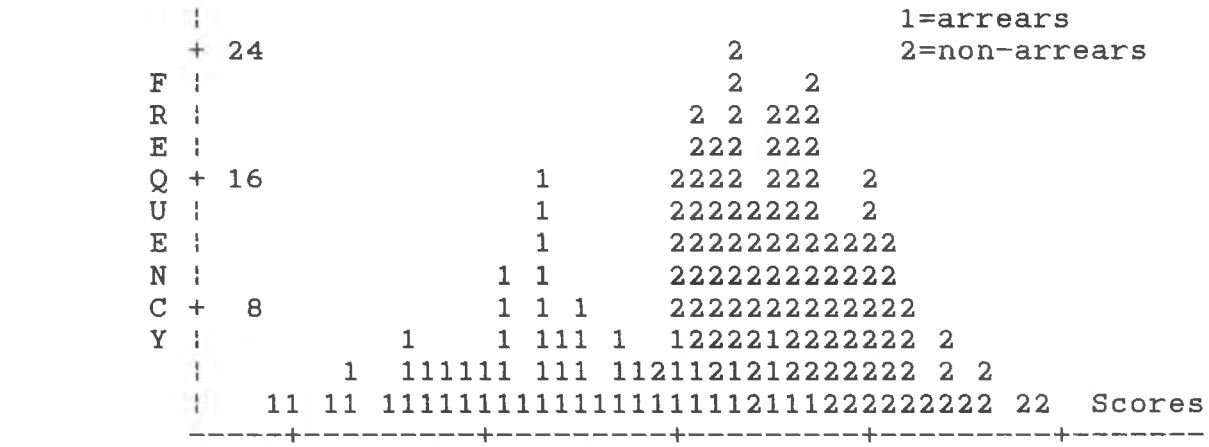
Estimated discriminant function	Partial F(5,296)	Mosteller-Wallace Contribution
z = 6.33843		
+ 0.07470 NARY (net assets ratio)	164.4	53.3%
- 0.06703 DCPI (inflation)	47.5	21.0%
- 0.06836 INPS (int./debt service)	19.2	13.9%
+ 0.12368 DGDP (GDP growth rate)	8.1	6.3%
+ 0.07434 INVR (investment/GDP)	6.5	5.5%
		100.0%

Wilks's lambda = 0.41 Box's M = 341.7
Chi squared(5) = 268.2 Approx. F(15,172213)=22.3

The best fitting discriminant function included five variables, and is set out in Table VII.1. It turns out that the variables in the model represent the first five principal components of the data set. They are listed in their order of entry into the discriminant function. The partial F ratios ('F-to-remove'), which are all significant at the 5 per cent level, reveal the same ranking and give an indication of the relative importance of the discriminating variables. The chi-squared value associated with Wilks's lambda provides a test of the hypothesis that the model has no discriminating power, and this hypothesis is rejected at the 5 per cent level.

Figure VII.1: Discriminant model I: all-groups stacked histogram of discriminant scores

---



Finally, visual inspection of the all-groups stacked

histogram of discriminant scores suggests that the estimated discriminant function yields a good separation of the two groups.

VII.3.4 Testing multivariate normality

Two maintained hypotheses of the linear discriminant model are: multivariate normality of the predictors within each group, and equality of the within-group covariance matrices. Of the variables in model I, four are approximately marginally normally distributed, in that on the basis of either the Kolmogorov-Smirnov or Jarque-Bera tests (see Appendix VII.1), the hypothesis of univariate normality cannot be rejected at the five percent level. The fifth and exceptional case is DCPI, the growth rate of consumer prices. For it, univariate normality is rejected at the five percent level, and indeed at the one percent level. Other confirmation of these findings is provided by the normal probability plots, which deviate greatly from linearity only in the case of DCPI, and also from the values of the optimal power transforms (see discussion in chapter IV). The optimal values for DCPI are 0.1 and 0.4 for the arrears and non-arrears groups respectively, indicating deviations from normality in each case. The optimal values of the power transforms for the

other variables all lie in the range 1.0 to 1.3, indicating a much closer approach to normality, except for the variable INVR (investment/GDP) in the arrears group, where the value is 0.5.

Koziol (1982) has developed a test for multivariate normality of a vector variable x. It is based on the quadratic form

$$(x-u)'V\,(x-u)$$

where u is the vector of population means, and V is the inverse of the population covariance matrix of x. Under the null hypothesis of multivariate normality, this quadratic form has a chi-squared distribution with degrees of freedom equal to the number of variables. To apply Koziol's test, the sample of size n is used to estimate V and u. Clearly, the sample itself consists of a set of vectors x(i), i=1,n, drawn from the underlying population, and these together with the estimates of V and u are used to calculate a test statistic. Details of this calculation are given in Koziol (1982), along with a table of critical values. A recent application of the test may be found in Watson (1990), where it is applied to a set of corporate financial ratios.

Koziol's test is applied here to the set of five discriminating variables that are used in model I, the test being applied separately to the arrears

group (n=102) and the non-arrears group (n=200), the computed values of the test statistic being 0.0584 and 0.1439 respectively. In both cases, degrees of freedom = 5. These parameters must be approximated in entering Koziol's table, where critical values of the test statistic are given for 2 and 6 d.f., and for n = 100 and for n tending to infinity. Where K is the computed value of the test statistic, we have:

for 2 d.f. and n=100: $P(K < 0.2060 \mid Ho) = 0.95$
for 6 d.f. and n=100: $P(K < 0.2011 \mid Ho) = 0.95$

while for n tending to infinity:

for 2 d.f. : $P(K < 0.2224 \mid Ho) = 0.95$
for 6 d.f. : $P(K < 0.1852 \mid Ho) = 0.95$

The computed values of K are below the appropriate critical values in all cases, and therefore we cannot reject the null hypothesis (i.e. multivariate normality of the five variables) at the five per cent significance level.

This result is at odds with the univariate analysis, where the hypothesis of normality, which is a necessary condition for multivariate normality, is rejected for the variable DCPI. There are two possibilities here. One is that the data are not multivariate normal, and that Koziol's test has generated a type II error. Alternatively, it may be that the extent of non-normality of DCPI is offset by the presence in the data of four other variables that are approximately univariate normal.

In chapter IV.6.2, a case was made, on several grounds, against working with transformed variables. If the result of Koziol's test be accepted at face value, then the implication is that it is unnecessary to consider transformations of the five-variable subset that is used in model I. However, an interactive stepwise discriminant analysis was carried out on the transformed data, where each variable had been subjected to its optimal power transform (chapter IV.6.1). The optimal discriminant model that emerged from this turned out to contain the same five variables as model I, after applying power transforms using the exponents that are set out in Table VIIA.1. Moreover, its classificatory performance was identical. This finding suggests that the specification of model I is robust.

VII.3.5 Tests of equality of dispersion matrices

The second maintained hypothesis, of equality of the dispersion matrices, is tested by Box's M-test. This comes second, because the test itself is sensitive to departures from normality (Mardia, 1974). Given the value of the approximate F statistic associated with Box's M (see Table VII.1), the hypothesis of equality is rejected at the 5 per cent level. In principal this indicates that the linear specification should

be rejected in favour of a quadratic discriminant function. This course has not been followed, because applications of quadratic discriminant analysis have not generally produced good results. According to Frydman et al (1985, p.278):

> Quadratic models, however, while quite accurate on original samples, are uniformly disappointing vis-a-vis linear models in holdout sample tests.

VII.3.6 Interpretation of model I

The discriminant function classifies as 'non arrears' those cases whose z-score exceeds a cutoff value. As would be expected a priori, the signs of the coefficients imply that a non-arrears classification is associated with high values of net assets/GDP, GDP growth and fixed investment/GDP, and with low values of inflation and of interest/debt service.

The relative importance of each of these variables is indicated by the Mosteller-Wallace contribution, as set out in Table VII.1. This reveals the share of each variable in accounting for the Mahalanobis distance between the group centroids; the partial F-ratios, in contrast, indicate the rank order, but not the precise shares. On this basis, the net assets ratio is clearly the major source of

discriminating power, with inflation and interest/debt service also important. The remaining two variables have a combined contribution of only 12 per cent. The discriminant function was re-computed using only the first three variables, but this model had a higher value (43.5) of Wilks's lambda, and a less distinct separation between the two groups, as revealed by the all-groups stacked histogram. On this basis, the five-variable model was retained.

A feature of this model is the absence of any type of debt-service ratio, although several such ratios are included in the data set, including:

PDSPX Projected public debt service/exports
TDSR (Total debt service + short term debt)/exports
TSPX Total external debt service/exports.

A related ratio is
TSPY Total debt service/GDP.

Within the set of five explanatory variables, the debt and debt-service dimensions are represented by the net assets ratio NARY, and the variable INPS which is concerned with the structure of debt service rather than its level. However, the absence of a debt-service ratio conforms with the findings of other recent studies of country risk. For example, the debt service ratio appears as an explanatory ratio in the final models of Feder and Just (1980) Feder et al (1981), Kugler (1984), and Edwards

(1984). However, Feder and Just found the ratio not to be significant at the five per cent level, while for Kugler it had the wrong sign in one model and was not significant otherwise. In Edwards's paper, the dependent variable was spread over LIBOR, which he used as an indicator of default risk. In a logit analysis, the debt-service ratio had a significant coefficient at the five per cent level in only one of six models estimated. By comparison, in none of the following papers had the ratio any role: Taffler and Abassi (1984), Kharas (1984), Schmidt (1984), Morgan (1986), Berg et al (1988), Lloyd-Ellis et al (1989).

Burton and Inoue (1985a) have written:

> Not surprisingly...empirical studies find little significance in this ratio.

There are four possible reasons for this.

(1) Debt service ratios measure debt service paid, not owing. A high level of debt service owing may be a leading indicator of servicing difficulties, but the measured ratio will by definition fall as arrears start to accrue, giving a perverse signal unless the timing of the onset of arrears can be established very accurately.

(2) Countries are typically faced with peaks and

troughs in service commitments. A country may be currently in a trough, and not in arrears, and yet be approaching a troublesome peak that will not be reflected in the current debt-service ratio.

(3) Debt service ratios may fall in consequence of debt restructurings, and thus give misleading indications about countries that are still in difficulties.

(4) A country may have a low debt service ratio because, through being perceived as not creditworthy, it has been unable to borrow.

In an attempt to overcome the difficulties with the ex post debt service ratio, the data set contains a variable PDSPX that is the ratio to exports of estimated debt service, one year ahead. A simple test established that this variable is a more effective discriminator than any of the ratios

TSPX (total debt service/exports)
TSPY (total debt service/GDP)
TDSR (total debt service + short term debt)/exports)

For each of the four ratios under review, a linear discriminant function was estimated. For PDSPX, the value of Wilks's lambda was 0.651, whereas for the other three its value was, respectively, 0.839, 0.832 and 0.775. This superiority of PDSPX to the other

debt service ratios is explicable on the basis that it avoids the first of the four difficulties that are associated with these ratios. However, the other three remain, so that it is not surprising that models containing this variable had an inferior performance to model I, on the key issue of classificatory performance.

VII.3.7 Within sample classification: model I

The classificatory power of the model may be tested on the estimating sample. However, such a test is biased towards over-optimism, and to remove or reduce this bias the Lachenbruch holdout method ('U-test') may be used. (Lachenbruch and Mickey, 1968). Given the sample of size 302, 302 different sub-samples are formed of size 301. Each consists of the full sample less one case held out. In each case, a discriminant function is estimated on the sub sample, and then used to classify the held-out case.

On this basis, a set of discriminant scores is computed, some of which are set out in Table VII.2, while a summary of the holdout classifications is set out in Table VII.3.

Table VII.2: Discriminant scores: model I, 1979-86

Case				Scores (in numerical order)	
194	TW	86		9.31	
193	TT	80		8.98	
Section includes 175 non-arrears cases, 12 arrears:					
117	NG	85	*	104 DR 83 * 87 PK 80 * 86 DR 81 *	
84	GA	85	*	76 NG 82 * 74 LB 79 * 48 VZ 82 *	
41	YU	82	*	39 SE 80 * 24 MX 81 * 22 UR 82 *	
5	HD	80	*	0.09	First col. is order of
4	KN	83		0.08	case from cutoff (a)
3	DR	84	*	0.02	
2	EC	82	*	0.01	* denotes type I error
1	BD	83		0.00	relative to cutoff (a)
Original cutoff (a)					
1	CB	82	+	-0.05	
2	ES	86	+	-0.09	
3	VZ	86		-0.17	
4	EG	79	+	-0.20	+ denotes type II error
5	KN	85	+	-0.21	relative to cutoff (a)
6	BD	82	+	-0.28	
7	PP	84	+	-0.29	
8	ZI	84	+	-0.30	
9	PP	85	+	-0.35	
10	CB	85	+	-0.37	
11	ES	82	+	-0.41	
12	SR	83	+	-0.44	
13	BD	80	+	-0.45	
14	KN	82	+	-0.47	
15	ES	80	+	-0.47	
16	TK	86	+	-0.49.. (b)	revised cutoff
17	PH	82		-0.51	
18	MC	82		-0.54	
19	SR	81	+	-0.63	
20	SU	82		-0.75	
21	TK	81		-0.76	
22	EG	80	+	-0.76	
23	GU	86		-0.77	
24	TN	86	+	-0.81	
25	ES	81	+	-0.82	
26	BL	79		-0.84	
27	SE	81		-1.05	
28	KS	80	+	-1.27	
Below this point are 80 cases, all arrears.					

Note:
Date of predictor data is shown for cases: this is used to classify for following year.

Table VII.3: Discriminant model I: holdout classifications, original cutoff

		Predicted group membership 1980-1987	
Actual group:	Total	Arrears	Non-arrears
Arrears	102	87 (85.3%)	15 (14.7% type I errors)
Non arrears	200	20 (10.0% type II errors)	180 (90.0%)

The error rates are low, and the all-groups stacked histogram (Fig. VII.1) reveals clear separation of the two groups. Further inspection of Table VII.2 reveals that of the 20 type II errors, 2 lie beside the cutoff, followed by an arrears case, followed by a further 13 clustered type II errors, with the remaining 5 a little further away.

The optimal cutoff between the two groups is that which minimises the expected cost of misclassification. This is given (Mensah, 1983: p.23) by:

$$\text{Expected cost} = [C \times P \times T1/N1] + [(1-P) \times T2/N2]$$

where:
P : prior probability of 'arrears';
T1, T2: numbers of type I and II errors;
N1, N2: sample numbers of cases of each type;
C : cost of a type I error
(cost of a type 2 error = 1.0).

The ratios T1/N1 and T2/N2 are the conditional probabilities of type I and II errors respectively. However, if we assume that prior probabilities equal sample proportions, then minimising this expected cost is equivalent to minimising the ex post cost:

$$[C \times T1] + T2$$

The cutoff that minimises the total number of misclassifications is (b) in Table VII.2, with 16 of type I and 5 of type II. Where each type has the same cost, this is the ex post cost-minimising cutoff. Moreover, if type I misclassifications are given more weight, then for a wide range of weightings, (b) is the cost-minimising cutoff. This is the case if the cost of a type II error is set at 1.0 and the cost of a type I has any value in the interval [1.0, 4.2], which includes the original value of 3.75. Compared with cutoff (a) and Table VII.3, cutoff (b) reduces the number of type II errors to 5 (2.5 per cent), while raising the type I error rate by one case (VZ 87) to 16 (15.7 per cent). Table VII.4 shows the holdout classifications thus determined.

Compared with Table VII.3, the cutoff has been moved by 16 places, i.e. from (a) to (b) in Table VII.2, and the numbering of the cases in Table VII.4 reflects this: for example, Nigeria (dated 1985 for data, 1986 for debt-servicing status) was 117 places

from the original cutoff, and is 133 from the revised cutoff.

Table VII.4: Holdout classifications by discriminant model I. Intercept (cutoff) amended

Actual group:	Total	Predicted group membership 1980-1987: Arrears	Non-arrears
Arrears	102	86 (84.3%)	16 (15.7% type I errors)
Non arrears	200	5 (2.5% type II errors)	195 (97.5%)

Type II errors			Type I errors		
South Korea	1981	(12)	Nigeria	1986	(133)
El Salvador	1982	(9)	Dominican Republic	1984	(120)
Tunisia	1987	(8)	Pakistan	1981	(103)
Egypt	1981	(6)	Dominican Republic	1982	(102)
Sri Lanka	1982	(3)	Gabon	1986	(100)
			Nigeria	1983	(92)
Total 5 (2.5 per cent)			Liberia	1980	(90)
			Venezuela	1983	(64)
The numbers in parenthesis show the order of cases from the cutoff: the lower the number, the closer is the case to its correct group.			Yugoslavia	1983	(57)
			Senegal	1981	(55)
			Mexico	1982	(40)
			Uruguay	1983	(38)
			Honduras	1981	(21)
			Dominican Republic	1985	(19)
			Ecuador	1983	(18)
			Venezuela	1987	(14)
			Total: 16 (15.7 per cent)		

Notes: This table uses cutoff (b) of Table VII.2. The score for a country at date t is used to classify it at date t+1.

Frank, Massy and Morrison (1965) (hereinafter FMM), and also Morrison (1969), have suggested a nonparametric test of the null hypothesis that

classification by the discriminant function is no better than random. Under the null hypothesis the test statistic, which is derived from the classification matrix, has the t-distribution with degrees of freedom equal to sample size. The value of the test statistic in the present case is 13.2. This exceeds t(0.025, 302), so at the five per cent level (two-sided test), the null must be rejected in favour of the alternative hypothesis that the classificatory power of the discriminant function is better than random.

It is clear that while the type II errors are all close to the cutoff, this is not so for the type I cases. In this sense, the model is better at classifying the non-arrears cases, although the type I error rate is not unacceptably high. The type II errors were examined for evidence that the model gives warning of arrears with a lag longer than one year, i.e. advance early warning. In fact, four of these cases have never been in arrears since the date at which they were a type II error. The exception is Egypt, which went into arrears during the 1980s, and this provides little evidence for a re-interpretation of the type II errors.

VII.3.8 Model I: jackknifed coefficients

For out-of-sample prediction, and for the tests of the model to be reported in chapter XII, unbiased coefficients are required. Jackknifing is a general method that removes bias of the order [1.0/(sample size)] in the estimator of a parameter, and it is employed here. (Miller, 1974; Crask and Perreault, 1977; Taffler, 1984a). These jackknifed coefficients are set out in Table VII.5, along with the raw coefficients and the average pseudo-values from which they are computed. In computing the second and third columns, double precision arithmetic was used, before rounding.

Table VII.5: Coefficients for model I

	Raw coefficient	Average pseudo-value	Jackknifed coefficient
Interc.	6.83843	6.83898	6.67148
NARY	0.07470	0.07471	0.07319
DCPI	-0.06703	-0.06704	-0.06548
INPS	-0.06836	-0.06837	-0.06693
DGDP	0.12368	0.12369	0.12150
INVR	0.07434	0.07434	0.07327

Note:
The intercept is based on the amended value of the likelihood ratio (1.158)

VII.4 The data set augmented: categorical variables

VII.4.1 Introduction

The approach of this chapter is essentially two-stage. Section VII.3 sought an optimal discriminant function based only on continuous variables, given that only variables of that type should, strictly, be used in the linear discriminant model. However, in a number of dimensions, the sample may be partitioned into distinct groups. In the analysis so far, such partitionings have been ignored. Implicitly, it has been assumed that a given vector of values for the discriminating variables would have implications for debt-servicing capacity that would not depend on the country-year-case's group membership. This may be inappropriate: because of the diversity of the sample membership, it may be that the implications of a given financial and economic profile for creditworthiness may depend on group membership. This possibility is now addressed, by searching for categorical variables that have significant discriminating power.

This second stage thus remedies a possible

shortcoming of the original approach, at the cost of certainly violating one of the maintained hypotheses of the linear discriminant model: a set of variables that includes categorical variables cannot have a multivariate normal distribution. However, Krzanowski (1975) and Press and Wilson (1978) have found that the linear discriminant model retains its classifying power when the predictors include both continuous and binary variables. A retention of discriminating power under an even further departure from multivariate normality has been established by Amemiya and Powell (1983), in a study in which all the independent variables were binary.

The first possible partitioning is intertemporal - a limiting case would be where each year required separate treatment, while the simplest partition involves a binary split. The second possibility utilises the 'environmental' variables described in chapter IV.3.6, Table IV.4. Since each of these has a constant value across countries within a given year, these variables all involve an intertemporal partition of the sample. The third possibility involves the categorical variables listed in chapter IV.4, Table IV.5, including geographical region, oil/nonoil, heavy indebtedness, type of export, borrower status, and low income.

A fourth possibility is to utilise the grouping determined by membership of cluster I (see chapter VI), but this has been rejected. The eighteen countries in cluster I were all 'non-arrears' in the sample period, except India in 1980. They thus include 125 of the non-arrears cases. There would be little predictive power in a model that used as a predictor a variable that was so closely related to the variable to be predicted.

In most of the following, the discriminant model is augmented by the addition of binary-valued variables to the data set. For each type of classification, k-1 binary variables are defined, where there are k separate groups within the classification. To avoid multicollinearity, one group has no binary variable, and when a case belongs to it, all the binary variables equal zero. Otherwise, any given case will have zero values for all the binary variables except for the one corresponding to the group to which it belongs, which will equal unity. The hypothesis to be tested is that group membership affects the intercept of the discriminant function. In effect, in each of the cases discussed below, the null hypothesis is that a given binary (dummy) variable has no discriminating power, and this is explored using an interactive stepwise discriminant analysis. The data set is the primary set of 70 financial and

economic variables, augmented by the appropriate dummies. As described below, in most cases the null hypothesis could not be rejected - typically, the first five variables to be chosen were those listed in Table VII.1, while the F-to-enter statistics of the dummies were well below 4.0.

VII.4.2 Intertemporal partitioning

In the recent history of external debt problems, 1982 was a significant year - as chapter I makes clear, that was when the crisis broke. It is therefore appropriate to partition the sample at or about 1982. Attempts were made to find a discriminant function containing a significant binary variable with a value 1 in or after 1982 and 0 otherwise. Alternative cutoff dates of 1981 and 1983 were also examined. In conclusion, the null hypothesis could not be rejected.

An alternative approach to 'time' involves re-estimating the discriminant function on pre- and post-1982 sub-samples. Since this raises questions concerning the stability of the specification of the function, it will be dealt with in chapter X.

VII.4.3 'Environmental' partitioning

Dummy variables were defined with reference to the external, environmental variables - these include the growth rates of world trade, of GNP in the industrial countries, and of import and export unit values; fiscal balance/GNP in the G7 countries; and nominal and real US$LIBOR. In no case could the null hypothesis be rejected.

VII.4.4 Other classifications: model II

In the case of the following categorical variables, the null hypothesis could not be rejected:

KMX: official borrower versus other categories
LIC: low income versus other countries
OIL: non-oil versus other countries
PRX: predominant export:
-fuel
-mineral
-agriculture,
-manufacturing;
-service and remittance.
REG: Africa, Western Hemisphere.

The two cases where the null was rejected were: REG = Asia (versus all other cases), and HID = heavily indebted country (versus all other cases). However, the HID variable was unsatisfactory for two reasons. First, it is open to similar objections to those raised against using membership of cluster I: it

picks a group of countries that are known a priori to account for 46 of 102 arrears cases within the sample. Secondly, when a principal components analysis was run on the primary data set with this variable added, it emerged that it and NARY were loaded on the same component, with loadings of 0.46 and 0.56 respectively. Since NARY entered first into the discriminant function, HID was eliminated.

The Asian dummy entered the equation at the fourth step, after NARY, DCPI and INPS. Further steps made no significant improvement to Wilks's lambda or to the classificatory power of the model, so the process was terminated. The summary output is set out in Table VII.6. The partial F-ratios and Wilks's lambda are all significant at the 5 per cent level. The histogram of Figure VII.2 reveals clear separation of the two groups.

The intercept shown in the table is based on a cost analysis, as applied previously to model I. It yields cost-minimising holdout classifications for all values of costs of type I errors in the interval [2.0, 6.0], with the cost of type II errors set at 1.0.

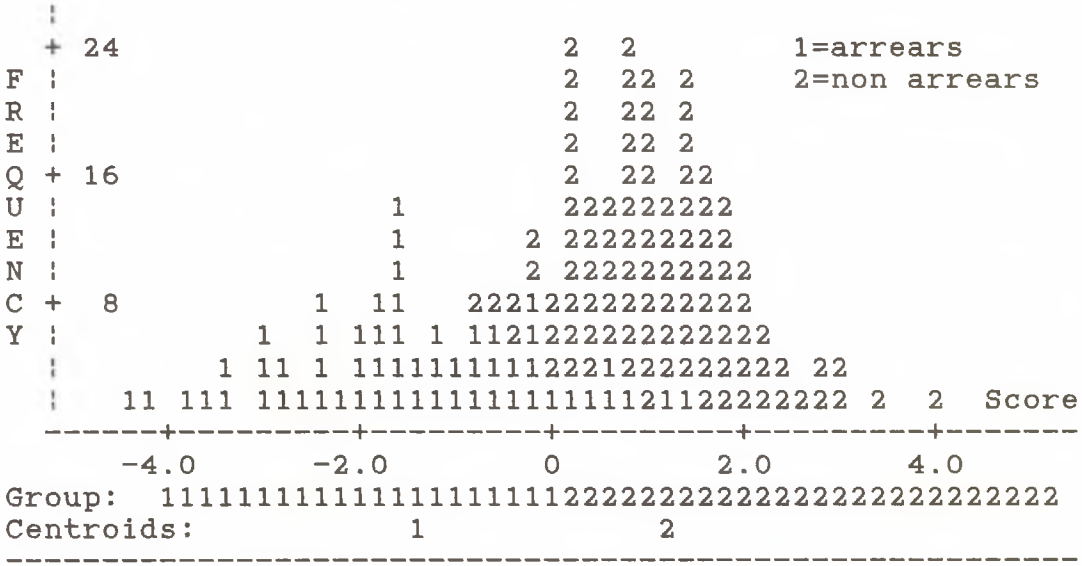
This model and its predecessor will be referred to as discriminant models I and II, respectively.

Table VII.6: Discriminant model II

Estimated discriminant function		Partial F(1,300)	Mosteller-Wallace contribution
z = 8.10520			
+ 0.07128	NARY (net assets ratio)	194.6	51.6%
- 0.06433	DCPI (inflation)	75.4	20.5%
- 0.08006	INPS (int./debt service)	78.8	16.4%
+ 2.00772	AS (=1 if Asia, else 0)	53.4	11.5%
			100.0%

Wilks's lambda = 0.41 Box's M = 396.4
Chi squared(4) = 266.5 Approx. F(10,198646)=39.0

Figure VII.2: Discriminant model II: all-groups stacked histogram of discriminant scores



In terms of the first three variables, models I and II are very similar: they have similar coefficients, with similar significance levels, and similar Mosteller-Wallace contributions.

In model I, we could not reject the hypothesis of multivariate normality. A priori, we should not expect the four variables in model II to be multivariate normal, given the inclusion of the categorical variable AS. The values of the Koziol statistic (see section VII.3.4) for the sample values of these four variables are 0.981 and 0.233 for the arrears and non-arrears groups respectively. Comparing these with the critical values, the (null) hypothesis of multivariate normality is rejected at the five percent level in each case. Moreover, given the value for the F statistic associated with Box's M (see Table VII.6), the (null) hypothesis of equality between the within-group dispersion matrices is rejected. However, this latter result should be treated with caution, in view of the apparent non-normality of the data (Mardia, 1974).

The holdout discriminant scores of the two models are highly correlated: the correlation coefficient is 0.96, and is significant at the five percent level. However, the classifications differ somewhat; those of model II are set out in Table VII.7. The AS Asia variable in model II has a positive coefficient, implying that membership of this group, for given values of the other discriminating variables, makes it more likely that a case will be classified as non-arrears. This confirms a priori expectations.

Table VII.7: Holdout classifications by model II

Actual group:	Total	Predicted group membership 1980-1987: Arrears	Non-arrears
Arrears	102	91 (89.2%)	11 (10.8% type I errors)
Non-arrears	200	20 (10.0% type II errors)	180 (90.0%)

Type II

Congo	1982	(52)
Congo	1981	(42)
Egypt	1981	(39)*
Egypt	1980	(35)
Turkey	1987	(32)
Congo	1980	(29)
Egypt	1982	(25)
Tunisia	1986	(23)
Tunisia	1987	(22)*
Tunisia	1983	(21)
Tunisia	1985	(20)
Colombia	1986	(17)
Tunisia	1984	(14)
Kenya	1986	(11)
Turkey	1986	(10)
Colombia	1983	(9)
Kenya	1983	(5)
S. Lanka	1982	(5)*
Jordan	1986	(3)
Kenya	1984	(2)

Total: 20, 10.0 per cent

Type I

Nigeria	1986	(131)*
Pakistan	1981	(123)*
Dominican Repub.	1984	(75)*
Nigeria	1983	(57)*
Gabon	1986	(55)*
Uruguay	1983	(51)*
Senegal	1981	(50)*
Dominican Repub.	1982	(42)*
Liberia	1980	(31)*
Philippines	1983	(25)
Venezuela	1983	(20)*

Total: 11, 10.8 per cent

Notes: Numbers in parenthesis: see Table VII.4
*: indicates that case was misclassified by model I

It may also explain why model II generates many more type II error cases than model I (20 versus 10), given that only one of these is an Asian country. The model avoids the misclassification made by model

I of Korea (1981), but not that of Sri Lanka (1982). Evidently it can identify non-arrears cases in Asia but is not as efficient at doing so elsewhere. Only three type II error cases appear in both models, whereas the overlap between them is much closer for the type I cases, with 10 in common.

Compared with model I, there are fewer type I errors (11 versus 16). The computed value of the FMM statistic is 12.2, so as in the case of model I the null hypothesis, that overall classification by the model is no better than random, is rejected at the five per cent level.

Ultimately the choice between the models must be determined by out-of-sample performance. This will be addressed in chapter XI; jackknifed coefficients will be required, and these are set out in Table VII.8.

Table VII.8: Coefficients for model II

	Raw coeff-icient	Average pseudo-value	Jackknifed coeff-icient
Interc.	8.10520	8.10572	7.94631
NARY	0.07128	0.07128	0.06997
DCPI	-0.06433	-0.06433	-0.06310
INPS	-0.08006	-0.08006	-0.07896
ASIA	2.00772	2.00778	1.99158

VII.5 The problem of weak-year cases

VII.5.1 Multiple discriminant analysis

Thus far, this chapter has utilised two-group discriminant analysis, with the weak-year cases excluded from the data. However, this will present a problem when applying the model to out-of-sample data, and assessing its performance in that context, because some of these data will be weak-year cases. Multiple discriminant analysis (MDA) is an extension of the two-group model to circumstances in which the data fall into three or more groups. An attempt has been made to develop a MDA model, using the full 440-case sample. Two approaches were used. First, all the weak-year cases were formed into a third group. This group consisted of country-year cases for which ARS(t+1) = 1,2,3 or 4. Alternatively, those having the categorical variable ARS(t+1) coded 1 were defined to form a third group, while the other weak-year cases were defined to form a fourth group. The rationale of the latter approach is that weak-year cases themselves are not homogeneous, and that a distinction may be drawn between those that are closer to an arrears year, and those that are further away. In each case, an interactive stepwise analysis was run.

Details of these analyses are set out in Appendix VII.2. In summary, the variables that appeared in the optimal discriminant functions were very similar to those in model I. In the first case the only difference was that the fourth principal component (see chapter V) was represented by RYPC (real income per capita) rather than INVR (fixed investment/GDP). In the second case, there were four variables only: component 3 (represented previously by INPS, interest/debt service) did not appear; component 4 was represented by OCTD (official creditors/total debt); and component 5 was represented by DIND (growth rate of industrial production) rather than DGDP (growth rate of GDP) as in the other models.

The classificatory performance of the models was tested by the Lachenbruch holdout method, with prior probabilities represented by sample proportions. Both of the MDA models performed very poorly for the weak-year cases. In the three-group model, only 34 of 138 weak-year cases were correctly classified. In the four-group model, none of 77 cases with ARS(t+1) = 1 had a correct classification, and only 9 of the other 61 weak-year cases were classified correctly. The within-sample holdout classifications are shown in Tables VII.9 and VII.10.

Table VII.9: Holdout classifications of 440-case sample by three-group MDA model

ARS is at date t+1	Cases classified as ARS(t+1)= 0	1,2,3 or 4	9
Actual group: ARS(t+1)=			
0	79	16	7
1,2,3 or 4	78	34	26
9	17	30	153

Table VII.10: Holdout classifications of 440-case sample by four-group MDA model

ARS is at date t+1	Cases classified as ARS(t+1)= 0	1	2,3 or 4	9
Actual group: ARS(t+1)=				
0	90	0	3	9
1	62	0	3	12
2,3 or 4	34	0	9	18
9	31	0	14	155

VII.5.2 Weak-year cases and the two-group models

Given the extremely poor performance of the MDA models, the MDA approach is not pursued further. The alternative is to deal with the weak-year cases within the framework of the original two-group models. Ideally, the ordering by discriminant scores would reflect the ordering by the arrears variable ARS on its 0 to 9 scale. However, when the

jackknifed model I of Table VII.5 is applied to the full (440-case) sample, no such ordering emerges, as Table VII.11 reveals. Fifty-eight per cent of cases with ARS=2 are classified as non-arrears. In proportionate terms, cases with ARS=3 are similarly classified, and consequently cases with ARS=2, 3 or 4 are shown grouped. (Only three cases have ARS=4.) A similar exercise for model II produced almost the same results.

Table VII.11: Classification of weak-year cases by model I: summary, and distribution of scores

CLASSIFIED BY MODEL: ARREARS / NON-ARREARS	Decile	Numbers of cases: Actual group 1	Numbers of cases: ARS(t+1) = 2, 3 or 4
ARREARS	1st	20	7
	2nd	11	2
	3rd	12	5
	4th	12	12
	Totals	55	26
	Cutoff		
NON-ARREARS	Totals	22	35
	5th	7	6
	6th	3	5
	7th	9	8
	8th	1	6
	9th	1	7
	10th	1	3

Notes:
(1) Deciles are of scores of 440-case sample.
(2) Cutoff of zero separates cases classified as arrears (negative scores) from those classified as non-arrears (positive scores).

Table VII.11 shows that the discriminant model allocates negative scores to a majority of cases with ARS=1, and positive scores to a majority of cases with ARS=2,3,4. However, a sizeable minority in each case do not follow this pattern: the model is not decisively assigning the two categories of weak-year case to the alternative groups. Nor is the model assigning weak-year cases of all types to one group: 81 weak-year cases are assigned negative scores, while the remaining 57 have positive scores.

In principle, this could be a consequence of the precise location of the cutoff that has been used: in this case, the final cutoff that was developed in section VII.3.7. In fact, it is clear from Table VII.11 that it is not possible to get a clearer classification by moving the cutoff: there is insufficient skewness in the sample distributions of the discriminant scores.

A final hypothesis is that the weak-year cases, being intermediate between the arrears and non-arrears cases in creditworthiness, should occupy an intermediate position in the range of discriminant scores. Again, the evidence does not support this. With approximately 44 cases of all types in each decile, it may be seen that in no case does either type of weak-year case (ARS=1, ARS=2,3,4) form a

majority of a decile, while only in the cases of deciles 1 and 4 do weak-year cases of all types form a majority. Decile 4 also contains a significant minority (10 cases) of arrears years, so it cannot be concluded that this range of discriminant scores is identifying weak-year cases. The 20 cases with ARS=1 in decile 1 constitute the largest concentration of that type of weak-year case. It is unsatisfactory that this extreme of the set of ordered scores has this feature.

The implication of these findings is that the two-group discriminant models cannot identify weak-year cases: it is not possible to partition the range of scores into different sections for arrears, weak-year and non-arrears cases.

VII.6 Summary and conclusions

In this chapter, two linear discriminant models have been specified, estimated, and subjected to statistical tests. Both models yield statistically significant discrimination between the groups of arrears and non-arrears cases. Their coefficients have the correct signs, and all the partial F-ratios are statistically significant. The question of multivariate normality, which was left open in chapter IV, has been addressed. Using Koziol's test, multivariate normality is rejected for model II (which contains a binary variable), but cannot be rejected for model I.

The holdout within-sample misclassification rates of the two models are quite low, and using the FMM nonparametric test, the hypothesis of no classificatory ability is rejected for both models.

Jackknifed coefficients, and least-cost cutoffs for classificatory purposes, have all been estimated and will be used in chapter XI.

Finally, it has been shown that the discriminant method cannot handle the weak-year cases, via models I or II, or using multiple discriminant analysis.

Appendix VII.1

Table VIIA.1: Tests for univariate normality

Arrears group	NARY	DCPI	INPS	DGDP	INVR
Jarque-Bera	5.9	26.9	1.2	0.2	2.1
Kolmogorov Smirnov	0.10	0.20	0.06	0.04	0.10
Optimal power trans.	1.1	0.1	1.3	1.0	0.5
Non-arrears group					
Jarque-Bera chi-sq.	1.4	17.7	1.0	1.0	0.9
Kolmogorov Smirnov	0.05	0.10	0.03	0.05	0.05
Optimal power trans.	1.2	0.4	1.0	1.1	1.1

Critical values (alpha=0.05):
Jarque-Bera test (1-sided): chi-square (2)=5.991

Kolmogorov-Smirnov (2-sided): 0.135 for N=102 (arrears) cases; 0.096 for N=200 (non-arrears) cases. (Conover, 1980: Table A14.)

Note
If the optimal power transform is 1.0, then the variable untransformed is the best approximation to normality, compared with any other power transform between -3 and +3.

Appendix VII.2

Multiple discriminant analysis

(a) Three groups defined with:

$$\begin{aligned} & ARS(t+1) = 0 \\ 0 < & ARS(t+1) < 9 \\ & ARS(t+1) = 9. \end{aligned}$$

These analyses were undertaken on the full 440-case sample, using the interactive stepwise technique. In order of entry, and with the number of the principal component of the data set on which each is highly loaded, these variables were:

NARY net assets ratio	1
DCPI growth rate, consumer prices	2
RYPC real income per capita	4
INPS interest/debt service	3
DGDP growth rate, real GDP	5

All partial F-ratios significant at 5 per cent level.

(b) Four groups defined with:

$$\begin{aligned} & ARS(t+1)=0 \\ & ARS(t+1)=1 \\ 1 < & ARS(t+1) < 9 \\ & ARS(t+1)=9. \end{aligned}$$

Again, these analyses were undertaken on the full 440-case sample, using the interactive stepwise technique.

The discriminating variables were:

NARY net assets ratio	1
DCPI growth rate, consumer prices	2
OCTD official creditors/total debt	4
(no variable represented component	3)
DIND growth rate, industrial output	5

All partial F-ratios significant at 5 per cent level.