IMPERFECT INFORMATION AND FINANCIAL LIBERALIZATION IN LDCS

MARY GERARDA ZEPHIRIN

APRIL 1990

Thesis submitted for the Degree of Doctor of Philosophy

University of Warwick

Department of Economics

CONTENTS

List of Tables

List of Figures

Summary

This thesis examines the interest rate, market entry and credit decisions which banks are expected to make following financial liberalization. It uses analytic tools from information economics and industrial organization theory to consider the policy implications of behaviour which responds to the constraints of imperfect information. The financial markets of the Caribbean Commonwealth supply the stylized facts which inform the analysis.

Chapter 1 introduces the topics treated.

The financial liberalization hypothesis is based on the 1973 works of McKinnon and Shaw. Chapter 2 describes their characterization of market fundamentals and behaviour in LDCs. It discusses the descriptions of equilibrium, and the welfare implications of these equilibria, in models which analyze economies with similar fundamentals.

Chapter 3 derives stylized facts from the descriptions of the economic institutions and financial systems of the four Caribbean countries whose banking behaviour we model.

Chapter 4 analyzes deposit rate determination by banks in the long-run equilibrium of a search market. It posits that in long-run equilibrium depositors find it costly to switch banks because doing so requires that they forego improved service at their current banks. The inelastic deposit supply which results from these switching costs implies that monopsonistic deposit rates are a noncooperative equilibrium. It is argued that this facilitates tacit collusion among banks and that a deposit rate floor is the appropriate policy corrective.

Chapter 5 argues that the enhancement of intermediation service responsible for depositor switching costs reflects the information banks acquire about customers and their ability to offer suitably tailored service.

Chapter 6 considers bank entry into a market where established customers of certain value have switching costs. Entering banks attract new customers of lower expected value. If new banks are therefore unable to generate sufficient revenue to cover their fixed costs, they exit. This chapter argues that liberal entry policy is not sufficient to ensure competition.

Chapter 7 develops a simple model of bank screening by loan size in one sector of an economy. It finds a sequential equilibrium in which low-risk borrowers self-select by the choice of contracts with a loan size below that they demand at the interest rate for their risk class.

In Chapter 8 the partial equilibrium model of Chapter 7 is embedded in a general equilibrium framework to demonstrate that the market equilibrium is not constrained Pareto efficient. It argues that subsidizing the highest interest rates will improve loan allocation while maintaining the separation induced by private contracts.

Chapter 9 summarizes the main results and conclusions of the thesis.

Chapter 1: Introduction

Financial liberalization has become a major plank of the economic adjustment programmes implemented in LDCs. It involves the removal of regulatory interventions in the financial sector, it being presumed that interest rates will then rise to levels which attract domestic savings to the financial sector, reflect the scarcity of capital and are competitive with the returns on foreign markets. The resulting expansion in funds available to financial intermediaries, at rates of return which compensate them for their costs, permits intermediaries to fund the investment projects which they judge economically viable, that is, finance is deepened. Analyses of financial deepening commonly focus on the relation between real monetary expansion and investment, using sectoral macroeconomic models to examine the results of liberalization. This thesis focuses on financial intermediary behaviour in economies with imperfect and costly information where agents are heterogeneous and untrustworthy. Our aim is to investigate three aspects of the post-liberalization financial intermediation described in the McKinnon and Shaw [1973] analyses. These are: deposit rate determination, the competitive forces provided by entry into the banking market, and credit allocation. We will also consider the policy implications of the behaviour derived in the first and third of these analyses. Observations from four Caribbean Commonwealth countries will form the basis of our environmental and institutional background. A complementary aim is therefore to account for observed outcomes in these economies with the aid of the hypotheses advanced by economists in the region.

In developing the thesis of financial deepening, McKinnon and Shaw, especially Shaw, were preoccupied with improving the allocation of resources in economies with segmented markets, heterogeneous labour and capital, a narrow span of financial assets and forward rates of interest, and expensive and incomplete information. Their advocacy of financial deepening was motivated by the role they attributed to the financial system whose richer information enables it to provide assets which can be traded intertemporally and interregionally, hence establishing prices which would reflect the relative availabilities of productive factors. This was the source of their stress on interest rates. Market determination of interest rates would unify segmented capi-

tal markets whose integration was seen as the key to the integration of the markets for labour and land. Financial regulation interfered with this process. In its absence deposit rates would increase to reflect scarce capital, competition would be improved by entry and banks would allocate credit to its best uses. We find, to the contrary, that when customer supply or demand is influenced by their incomplete knowledge of the market or by private information, collusive behaviour is sustainable, entry may be frustrated and credit is not allocated according to expected social return. Carefully directed policy can improve on the liberalized market outcome in our models, even though banks play the informational and transactional role ascribed to them by McKinnon and Shaw.

In a sense these results should not have surprised us. The expectation that the decentralized activity targetted by financial liberalization will produce a Pareto improvement is based on the welfare properties of the competitive general equilibrium. But, as Shaw can be interpreted as pointing out, the competitive equilibrium of standard analysis does not feature the heterogeneity, fragmentation and costly information common to lagging economies, and has no need for money and financial intermediation. The policy implications of our partial equilibrium models are conditional on their particular specifications. The wider point that public policy can almost always be Pareto improving in economies with the fundamental features described by McKinnon and Shaw must be based on more general models of market allocation. Chapter 2 is therefore a preamble where we look at the efficiency properties McKinnon and Shaw expected in a monetized and intermediated economy. The treatment of technical analysis in this chapter is essentially intuitive and lacks formal precision. It will suggest that the monetary mechanisms put forward by McKinnon and Shaw anticipate much of the more recent and formal analysis of markets with money and intermediaries. However, their work does not analyze the arrangements by which intermediaries acquire and use information, and hence precludes consideration of the welfare implications of these arrangements. Since they wrote, industrial organization theory and the economics of information have provided new insights into decision-making under imperfect information. It is these that we shall exploit to analyze deregulated behaviour.

Chapter 3 sketches the economic background and describes the financial sectors of the four Caribbean countries considered: The Bahamas, Barbados, Belize and Trinidad and Tobago. It describes the institutional framework and the observations we wish to explain, encapsulating these in stylized facts. It will be argued that one of the chief means by which banks acquire knowledge and enforce the behaviour of their clients is through the duration of relationships. This process creates an attachment between bank and client. Such attachments have recently been modelled in the industrial organization literature through the device of switching costs which capture an inelasticity of demand created by such attachment. Chapters 4 to 6 originate in the switching cost idea, extending it in the following directions. We begin by endogeneizing switching costs in the deposit market. This endogeneization suggests that switching costs arise from a gain to be made by remaining with a supplier (bank) and such gains are most important in a search environment where consumers have imperfect information about the product being purchased. Second, therefore, we will indicate the sources of that gain and describe some empirical representations of it in the banking market. The market significance of switching costs lies in the inelasticity of demand and sluggish market share which it creates. Other models have analyzed how these affect potential competition through entry prevention and cooperation between entrant and incumbents. Thirdly, we investigate a further aspect of potential competition, entry followed by exit, in markets with these characteristics. Chapters 7 and 8 go on to consider credit markets with asymmetric information.

Chapter 4 is concerned with explaining collusive deposit rate determination in the banking market. It will employ a dynamic model of savings, with search for suitable intermediation services, to derive the depositor decision. We shall show that a depositor who knows his current bank and expects improved service from remaining with it must balance that gain against the attraction of higher deposit rates elsewhere. This tradeoff creates switching costs in the long-run equilibrium of the banking market. In a mature market where all depositors have switching costs we will see that the noncooperative price equilibrium supports tacit price collusion, thus explaining the collusion hypothesis in the Caribbean banking market, as well as providing a possible explanation of noncompetitive outcomes in other mature markets. We will then show that the

effects of such tacit collusion can be corrected by interest rate regulation.

Chapter 5 is concerned with informally explaining how and why depositors can be confident of receiving enhanced service by remaining with their banks. It will be argued that the question can be viewed as an application of the contract duration reward arrangements most usually applied to the labour market. Their importance in the labour market arises from the asymmetry of information between employer and employee. But, we will argue, a similar asymmetry exists between bank and client: a client's prospects ('productivity') are both important to a bank and private information to the client. We will therefore argue that, by analogy with analyses of the advantages of long-term contracts, relationships between bank and client may be viewed as subject to an implicit contract. This contract is more easily enforced in the banking context because the award which rewards client longevity by an expansion in intertemporal trades is also of direct financial benefit to the bank. Such an implicit contract would explain the bank-customer relationship observed in banking markets. Further, the learning facilitated by, or embodied in, such contracts are an instance of the McKinnon and Shaw view of financial intermediaries as generators and users of information about relative values.

Chapter 6 will consider entry and exit in a market with the sluggish market shares associated with switching costs, using the observation of entry in Caribbean economies as its basis. It will argue that banks with imperfect information about prospective markets may enter them in search of new business even if aware that successful entry is not assured. We will show that the fewer the existing firms in such markets, the more likely is entry to succeed because the vested interests of incumbents encourage higher pricing and it is then easier for the new entrant to cover his higher costs. With a larger number of incumbents, it becomes more difficult for the entrant to attract known clients and, if market expansion proves of too little value, exit may be the entrant's optimal strategy. We interpret this entry-then-exit phenomenon as a search for information, rational in that the losses represent a cost of information. These results will furnish a further, to earlier chapters, instance of how switching costs confer benefits on longevity - in this case of suppliers - if analyzed in the context of imperfect information. They also explain similar phenomena in the history of the Caribbean banking markets.

Chapter 7 presents a model of bank screening in the credit market which demonstrates how information-constrained loan allocation diverges from expected social return. Credit rationing appears to be the only area of information economics and microeconomic theory that has previously been applied in the financial deepening context, probably because rationing appears an indisputable feature of so many LDC financial markets. This seems to contradict our argument that McKinnon's and Shaw's important insight was their characterisation of the financial system as an economical means of producing and utilising information, since rationing is symptomatic of banks' inability to distinguish between their heterogeneous clients. However, a variety of models which followed those with credit rationing have demonstrated that banks can offer sorting contracts which elicit information. In some of these, the means of sorting, although allowing perfect discrimination, may be observationally indistinguishable from rationing by loan size; or discrimination may be imperfect. Chapter 7 will discuss the application of the many models in this area to positive analysis in LDCs, and will develop a model which illustrates how some members of certain sectors can be receiving loans smaller than those demanded at the given rate of interest - although (in fact, because) banks are playing the kind of informational role discussed by McKinnon and Shaw. In addition to its welfare 'role' this model also employs a more appropriate equilibrium concept than is usual in screening analyses.

The model of Chapter 7 furnishes specific contract variables which we will use in Chapter 8 to assess the welfare implications of separating contracts in the loan market. Chapter 8 will initially investigate the policy recommendations which follow from models of rationing. As one would expect, a number of the models surveyed indicate that interest rate taxes and subsidies are appropriate. More surprisingly, we will employ an adaptation of the general equilibrium framework developed by Greenwald and Stiglitz [1986] to demonstrate that even when banks process information there remains a role for government subsidization of interest rates. In analyses of financial liberalization such credit subsidies are usually seen as one of the repressive government measures which impair financial deepening. This chapter suggests that they may be justified even when banks are efficiently playing their envisaged informational role.

Frequently, a thesis contains an introductory chapter which surveys the related literature. A different approach is adopted here: each chapter provides its own survey of relevant material; given the topics covered this approach seems to offer a more coherent treatment. Thus, with the exceptions of Chapters 2 and 3, each chapter discusses the relevant literature and, when necessary, places the subject considered in its Caribbean context before developing the model and discussing its results and implications.

The overall theme and conclusions of the thesis may be summarized as follows. Profit maximizing banks need to acquire information about their clients. However, they cannot trust them to convey the information truthfully when it is to the client's disadvantage, nor can the banks discriminate between liars and truthtellers. As hypothesized by McKinnon and Shaw, banks do have arrangements which help compensate for this imperfect information. Experience of clients is one means by which banks can assess reliability and they can write implicit service contracts with clients which enable them to acquire that experience. The attachments these contracts create seriously constrain competitive forces. Tacit collusion is facilitated. Further, potential entrants into the market can be prevented from acquiring a sufficient share of established customers to cover their costs. In the credit market banks have further means of acquiring information: screening contracts help discriminate among their customers. But these imply that marginal rates of substitution diverge. While these informational problems are common in financial markets, they are especially important in LDCs for at least two reasons. First, since banks are often the only sources of safe financial assets and external finance, neither savers nor borrowers can choose non-bank forms of screening. Second, growth potential in LDCs is probably greatest in innovative, non-traditional activities. When banks rely on experience as a source of information, those sectors are likely to be penalized. In these circumstances, financial liberalization may produce results not dissimilar from those it is intended to correct: low deposit rates, 'restricted' entry and rationed credit. Allocations are not in general constrained Pareto efficient. Consequently, there remains a role for regulatory measures and controls which are carefully adapted to the decentralized mechanisms which operate in liberalized financial systems.

Chapter 2: Financial Liberalization Theory from a Neoclassical Perspective

2.1: Introduction

This study will find that, with strategic decision-making under conditions of imperfect information, three propositions from McKinnon's and Shaw's financial liberalization hypotheses are not supported. Financial liberalization does not ensure that banks set deposit rates reflecting the opportunity cost of capital, liberal charter policy is not sufficient for competition-enhancing entry, and banks' comparative informational advantage does not permit them to allocate credit efficiently. These results are representative of the positive and welfare properties of the economies and economic organizations which form the basis of the McKinnon and Shaw analyses. Our specifications can only illustrate these properties. This chapter is therefore a preamble to consider the general analyses which derive these properties. Specifically, we discuss the informational role of prices, equilibrium in economies with incomplete markets and asymmetric information, and the money and financial intermediation mechanisms. The discussion will also, we hope, contribute to an appreciation of what we believe to be the key insight provided by McKinnon and Shaw: the financial system produces and transmits information which expands trading possibilities and efficient government policy must allow for this role.

The three propositions we will be examining were derived by McKinnon and Shaw from their general view that financial intermediaries play a crucial role in resource allocation and that that role is especially important in LDCs where information is imperfect, agents are heterogeneous and markets are incomplete. Intermediaries specialize in the production of information which allows them to set prices reflecting and signalling opportunity costs. These interest rates attract savings to banks who can allocate them efficiently as a result of their informational advantage. Removal of repression, where repression includes all governmental measures which tax or otherwise distort domestic capital markets[1], is seen as the key to allowing financial intermediaries to fulfill their role. Government's role should be confined to maintaining noninflationary mone-

---

[1] This summary description is given by McKinnon [1988]

tary growth and neutral fiscal policy.

We make three main points in this chapter. The first of these is positive. In economies with the fundamentals described by McKinnon and Shaw - asymmetric information, heterogeneity and incomplete markets - the Arrow-Debreu (AD) competitive equilibrium (CE) is not the appropriate equilibrium concept. The second point, also positive, concerns McKinnon's and Shaw's contributions. Recognizing that the processes assumed by neoclassical monetary growth theory models did not therefore apply to lagging economies, McKinnon and Shaw argued that in such economies the monetary system and financial intermediation played an irreplaceable role in transmitting information and completing markets. This assessment previewed both recent theories of general equilibrium with asymmetric information and money, and models of financial intermediation. These points would be of historical interest only if it were not for the third point which is about welfare implications. Given these functions of the monetary system, McKinnon and Shaw urged that government intervention should be minimized in order to allow financial intermediation to determine the prices required for information transmission and market completion. The economic rationale for minimal government intervention is the duality between AD CE and Pareto efficiency which assures us that in a CE with the AD conditions there is no government action which can make everyone better off. But we have already argued that the AD CE is not the appropriate equilibrium concept in the economies described by McKinnon and Shaw. That is, the fundamentals they described are inconsistent both with the existence and the Pareto efficiency of the AD CE. The welfare conclusions of the AD CE cannot therefore be applied in their economies. The appropriate equilibrium concept in such economies is now generally taken to be the rational expectations equilibrium (REE) or some financial assets CE. However the REE is not constrained Pareto efficient, and nor in general is any equilibrium with incomplete markets. That is, a central planner with no more information or markets than given by the CE could effect asset allocations or tax policies which make everyone better off. The existing structures are not efficiently used. McKinnon and Shaw's insights with regard to the informational and trade expansion roles of the monetary system were correct but, even after taking account of financial mechanisms, agents' marginal rates of substitution are not equated. This indicates that there are

government policies, in addition to macroeconomic stabilization, income redistribution and public good provision, that could in theory improve on the allocation of resources.

This brings us to the focus of the thesis. Many of the distortions described by McKinnon and Shaw, and attributed by them to official failure to allow the market system, especially finance, to fulfill its functions, are implied by the strategic reaction of agents to the fundamentals they described. That is, in economies with imperfect information, distortions such as price dispersion, monopoly power etc, would occur in the complete absence of government. Full liberalization would not therefore bring about an efficient allocation of resources. However, existing government regulations are almost certainly not those the central planner referred to above would implement. In effect, the systems before and after liberalization may not be Pareto comparable[2]. Liberalization may not bring about the expansion of savings, investment and employment suggested by the McKinnon-Shaw analysis. The lesson for policy is that liberalization is not sufficient. Existing ill-planned government intervention should be removed but policymakers then have the harder task of looking for appropriate taxes, subsidies and transfers to correct the distortions inherent in real economic systems.

The original McKinnon-Shaw policy recommendations were therefore incomplete. To explain why we believe this to be so this chapter describes their hypotheses from a microeconomic viewpoint and discusses the theoretical literature which rationalizes their views. Section 2.2 therefore highlights the environmental and behavioural background of their arguments, rather than the macroeconomic interpretation usually stressed. The following sections organize the discussion around three main themes emerging from our interpretation of their analyses. Section 2.3 considers the informational role of prices. In the McKinnon-Shaw theories it is the ability of interest rates to signal information about relative scarcities which permits an improved allocation of resources following liberalization. Section 2.4 discusses the literature on equilibrium when markets are incomplete, information is asymmetric and there is money, as in the McKinnon-Shaw economies. Section 2.5 surveys some of the recent literature which explains

---

[2] This is hard to accept since common sense suggests that the market must be able to improve on the pervasive but haphazard government regulations in many LDCs. The statement is therefore very tentative.

the functions and existence of financial intermediaries.

2.2: Policy Prescription and the McKinnon-Shaw Economies

Preliminaries

It is argued here that McKinnon and Shaw (MS) basically viewed financial intermediation as a mechanism, that is, as an organizational arrangement (specification of how) by which agents exchange information and make coordinated economic decisions. As such financial intermediaries help complete markets and thus reduce the effects of uncertainty: by expanding trading opportunities they permit individuals to insure against future uncertain events. Their determination of relative prices allows individuals to equate marginal rates of substitution in expectation terms. Financial intermediaries are able to play this role because they somehow have a comparative advantage in the production of information: one can view them as an analogue to the informed traders in the Grossman analysis (see Section 2.3 below). The financial system should therefore be liberalized to do its job, government pursuing only macroeconomic stabilization policy, public good provision and current budget surpluses on the fiscal front, while maintaining steady nominal money growth, setting a rediscount rate, ensuring easy entry and providing deposit insurance in the monetary sector[3].

In evaluating this proposition we may pose two questions. The first is whether the removal of existing repressive measures is sufficient to bring about an improvement in Pareto efficiency. This can only be answered on a country-specific basis since the answer depends on the regulations in place and the results hypothesized to follow their removal. The second is whether, from a situation where government is confined to the stance described above, there exists, at least in theory, government interventions which can bring about a Pareto improvement. The answer of this chapter is affirmative. Chapters 4 and 8 show in a very limited context two possible interventions. The answer is of current interest. Governments liberalizing their financial markets[4]

---

[3] The full package of measures for an unrepressed regime includes trade liberalization, tax-neutral fiscal policy, devaluation/slow depreciation to a free market exchange rate.

[4] We are of course abstracting from all the other problems that formed expectations and credibility are likely to create in this situation.

may not obtain the expected expansion of savings, investment and employment if the underlying market is subject to basic inefficiencies[5].

Our interpretation of McKinnon and Shaw may be questioned. Although these connections have not previously been made, as far as we are aware, in the LDC policy context, other analysts argue that market completion in the presence of imperfect information is the basic insight of McKinnon and Shaw-type financial analysis. For example, Bernanke and Gertler [1987] see recent financial intermediation theory as reviving the Gurley-Shaw view of financial systems and it is the earlier Gurley-Shaw views which MS extended and applied in LDCs (see the discussion in Ghatak [1981]). In addition, comparison of MS's work with contemporary analysis of distortion in the capital markets of LDCs highlights their focus on bank behaviour. Myint [1971], for example, attributed misallocation in capital markets to the financial dualism arising from unequal access to resources by the modern and traditional sectors; the unequal access itself being due to underdeveloped economic organization aggravated by government controls. He analyzed empirical outcomes similar to those addressed by MS: artificially low interest rates, chronic excess demand for loans and the resulting rationing of credit, and proposed increasing the official rate of interest in organized capital markets to reflect the shortage of capital funds (Myint, op. cit., p.331). Imperfect information, the signalling role of prices and the informational role of financial intermediaries play no part in his analysis. The following describes the MS hypotheses.

The McKinnon Hypothesis

McKinnon argued that the monetary sector reduced market imperfections due to heterogeneity, indivisibilities and imperfect information because money allowed banks to pool savings for investment and allocate these to high return uses. He therefore described money and capital as complementary.

Contrasting LDC environments with those featured in monetary growth models and neoclassical economies, McKinnon characterised LDC markets as " fragmented in the sense that firms

---

[5] As we shall see, especially in Chapter 8, the possibility of Pareto-improving government taxes and transfers indicates the presence of such inefficiencies.

and households ... face different effective prices for land, labour, capital and produced commodities and do not have access to the same technologies" (p.5)[6]. He attributes this fragmentation to three factors: two of these may be described as fundamental or "endemic in the underdeveloped environment without carefully considered public policy" (p.8), and a third arises from officially imposed constraints on the price mechanism.

Firstly, investment opportunities are indivisible because improved technologies require large discrete expenditures relative to individual resources: "without indivisibilities self-financed capital accumulation ...might well be sufficient for a slow diffusion of new technologies and a gradual reduction in the dispersion in rates of return within and between various enterprises" (p.13). Secondly, widespread uncertainty is generated by the heterogeneous information held by individual agents: "why 'subjective' uncertainty exists among firm-households is easy enough to see. ... Reliable information on any one contemplated loan or investment may be costly, relative to its size, for an outsider to obtain" (p.18).

These two fundamental features mean that real money balances and banking assume an increased importance. Money is risk- and default-free for short term transactions. It can be easily marketed, as compared to other financial instruments for which creditors require substantial information about the repayment capabilities of potential debtors. It is by frustrating the development of a financial market that the third factor of official intervention operates to maintain fragmentation.

As a result, capital endowments, productive opportunities and market opportunities for lending and borrowing are "badly correlated". The prevailing uncertainty prevents those with (discrete) investment opportunities from borrowing on the external market. They are therefore dependent on their own endowments. Government reacted to historical finance constraints by circumventing the domestic capital market through measures such as cheap credit and tariff protection. Such measures have distorted the allocative role of prices, with the result that the economy is fragmented: different agents face different prices which therefore fail to reflect/signal the

---

6 All page numbers in McKinnon and Shaw references refer to their [1973] books unless otherwise stated.

opportunity cost of resources.

In order to capture the process by which money and financial intermediation could relax the indivisibility and informational constraints in the absence of repressive measures, McKinnon devised a highly stylized model (Chapter 6) of an economy whose small agents are each both saver-investor and firm-household, confined to self-finance, with their small size implying that investment expenditure is "lumpy" relative to their resources. They are thus unable to undertake investments incorporating the most productive technology, and the resulting rates of return on physical capital are widely dispersed. In order to purchase physical capital other than their own output, agents must store own inventories for future sale, or accumulate cash balances. Money (where money is later defined to include both currency and the deposits of the banking system) is an important instrument of private capital accumulation in the uncertain environment because of its role as a means of payment and its sanction by the state[7]. Agents' willingness to hold it, and hence its value in this respect, depends on the rate of inflation, the nominal deposit rate and the 'convenience' of holding deposits and currency. The discrete rate of investment means that large purchasing power is required at the time of investment, so that the desired cash balances-to-income ratio rises with desired investment. It is in this sense that McKinnon argued that cash balances and real capital are complementary rather than substitutes as assumed in monetary growth models of industrialized countries.[8] Money in the lagging economies is a conduit for capital. It is only as the real deposit rate rises towards the marginal and intra-marginal return on self-financed investment that money and capital start to compete in the asset portfolio. This requires that the real return on money must, optimally, be less than the best return on self-financed investment in some units: if this is so the competing asset effect does not become dominant, discouraging real capital accumulation. Money's real return must also be positive, otherwise agents may invest in activities (e.g. storage) and goods whose internal rates of return are negative, but which may act as a store of value alternative to money holding. When there is a high return on cash bal-

---

[7] These two roles provide additional benefits from money holding and thus reduce its net holding costs relative to those of other instruments.

[8] Ghatak, op. cit., provides an apposite description of these models.

ances (low inflation) money will be accumulated for lump-sum investment.

Once explicit account is taken of financial institutions, the real return on money is optimally set higher because, rather than having to lie below the rates of return in some firm-households, it can reflect economy-wide rates of return as banks pool savings more efficiently. Full liberalization permits banks to attract savings and channel funds to all investors who can earn a high return, thus breaking the confines of self-finance. In order to extend credit bank loan officers must acquire costly information (p.77). It is these costs which, in a repressed economy, prevent banks from lending to small-scale investors, since regulatory ceilings do not allow loan rates to cover either the "peculiar administrative costs of serving each class of borrower" (p.79) or the default risks. Banks are thus prevented from competing with traditional lenders in the informal sector who acquire monopolistic power and charge correspondingly higher and dispersed lending rates (p.72).

Rate of interest regulations also constitute a barrier to entry in the financing of the traditional sector. Their removal would permit competition. Competition does not operate among the moneylenders because (in addition to legal restrictions) their detailed knowledge of only a narrow market restricts the scale of their operations, and because they lack the capital base for large-scale lending (p.78). It is presumed that banks will be able to compete effectively both because their pooling of savings overcomes capital constraints and because they have a comparative advantage in information acquisition.

The Shaw Hypothesis

Shaw's description of the lagging economy placed greater explicit stress on the informational role of prices and the place of banks (financial intermediaries) in implementing the market. Liberalization would "substitute a pricing mechanism and decentralized judgement" for the "rationing mechanism of repressed finance" (Shaw, p.133).

Shaw's description of the LDC environment, his debt-intermediation view (DIV), was set up to contrast with what he described as the wealth view (WV) of money and finance. Both views are macroeconomic but the underlying micro environment more or less explicit in his

analysis is as follows. The WV model exhibits many of the features expected in a Walrasian economy: trade is costless, there is perfect foresight and no uncertainty, agents and goods are homogeneous, commodities are divisible, information is costless and markets are complete. Shaw stresses that this model is inappropriate for the analysis of LDCs and that its use is at least partly responsible for government inflationary and repressive policies because, to put it briefly, WV ignores the role of money and finance (see, for example, p.102). The DIV is the antithesis of WV. Segmented markets, dispersed prices, heterogeneous and indivisible capital mean that savers and investors assess risks differently (p.50) and prices do not reflect social valuations. Extension of the monetary system and financial intermediation play a crucial role in compensating for these 'imperfections'. The financial system's services are an intermediate input into the production processes of others and could thus act to push out production possibility frontiers and extend markets (p.107). These services are attainable by encouraging the holding of money and deposits through decreases in expected inflation and a positive real deposit rate. Higher money balances would save resources in the search and bargain process (p.56), and permit prices to resume their function of clearing excess demands in interrelated markets (p.126), allowing the financial system to insure borrower and lender risks by providing a market where they can trade (p.127). The financial system has the ability to do this because it "presumably has clearer expectations and access to richer stores of information and can supply loan insurance" (p.128). Interest rates determined by banks would reflect intertemporal and spatial opportunity costs more accurately, discriminating between investment opportunities. Savers who lack the information to evaluate risky securities can delegate this task to the banks and spending units' ability to borrow would diminish the risk they face.

Finance uses real resources in allocating savings and facilitating money's role as a means of payment. The learning process is expensive (p.124). Differentials between loan and deposit rates must compensate the financial intermediaries for these costs. Unlike McKinnon, Shaw did not associate specialized information with monopoly power: the curb market's high rates reflect the risks associated with the illegality of their activities (p.89). With financial deepening it would become profitable for intermediaries to enter the market and "market forces" would push banks to

compete in the allocation of funds. Like McKinnon he viewed the absence of constraints as sufficient to stimulate competitive behaviour (pp.78-79).

Summary

There are differences in the analyses of McKinnon and Shaw. For example, McKinnon viewed money and capital as complementary because of indivisible technology and poor access to external finance. In Shaw money and capital remained substitutes but money's services joined capital in the economy's production function. It is their common features we stress. Both described economies with imperfect costly information, heterogeneous agents and goods and incomplete markets. Repression prevents the monetary system from fulfilling its pricing function so that prices are dispersed and do not reflect social costs, resources are misallocated and growth impeded.

Despite the role assigned to money and finance neither McKinnon nor Shaw investigate precisely how banks acquire information and make their pricing decisions, nor how they are able to pool savings efficiently. Behaviour and mechanisms are all summarized under market forces - which may be impeded by failure to compete but this is ensured by freedom of entry - and information transmission by prices. The superiority of market pricing is implied by the fundamental welfare theorems. However, the fundamentals and monetary arrangements which MS analyzed are not consistent with those in the economy of the welfare theorems. The processes they postulated are in part explained by analysis of the informational role of prices, and the role of money and financial institutions in the environmental conditions they described are derived in general equilibrium models with sequential trades (money) and in models of financial intermediation. But if we adopt these models we can no longer assume that official intervention is always inefficient. To indicate why this is so, the next three sections attempt to show how recent models can explain the general market mechanisms postulated by MS. The correspondence between these analyses and those of MS is obviously inexact. None of them for example take account of indivisible technology nor treat the case of large agents[9].

---

[9] That is, individual agents who are sufficiently large relative to the market to affect aggregates.

2.3: The Informational Role of Prices

We have argued above that the MS financial deepening argument rests on the informational role of freely determined interest rates. This reflects a widely-held view of the competitive price system discussed by Hayek [1945] but only rigorously examined by Grossman in the context of futures markets and stock markets in [1976] and [1977]. In the Arrow-Debreu (AD) economy with complete markets and state-dependent contingent contracts, prices have no role in transmitting and aggregating information (they do act to make actions consistent), since the uncertainty is aggregate: information is symmetric between agents. Informative prices require the use of rational expectations equilibrium (REE) prices where agents know the relation between prices and the stochastic determinants of output and use this information to make their decisions.

In order to capture the notion of the competitive price system as an economical mechanism for information transmission, Grossman found it necessary to replace Walrasian equilibrium (WE) prices by REE prices. In WE not only do agents ignore the information contained in prices[10], but observation of Walrasian market-clearing prices gives agents who make inferences from prices an incentive to recontract at WE prices[11], making it an inappropriate equilibrium concept (see Grossman [1981]).

Prices transmit information among insignificantly small agents if some invest in information collection; their information is reflected in their trades which influence price, transmitting the information to others (Grossman [1977]). Current price also aggregates information when agents have invested in different information since this information is reflected in their individual demands (Grossman [1976]). This transmission and aggregation is justified as the long run static equilibrium of a trading process: after repetition traders learn the joint distribution of the current price and the random future price. They are then able to condition their beliefs about future price on the current price and markets only clear at the current price which reflects all information, the REE price.

---

[10] Given the WE price an agent determining his trades looks only at his own preferences and budget constraint.

[11] Sophisticated traders learn that price provides information, on observing the equilibrium price therefore they will update their beliefs and wish to recontract.

The paradoxical nature of this equilibrium has been pointed out by Grossman, op.cit., and Grossman and Stiglitz [1980]. Once agents learn that the current price reflects all information (the REE is fully revealing), if investment in information is costly they have no incentive to acquire their own information since all necessary information can be costlessly learnt from observation of the current price. Since each agent considers his trading insignificant relative to the market, each will withdraw from information acquisition and the market price conveys no information. This paradox is most usually resolved by the assumption that prices are only partially revealing. This occurs if the current price is also affected by "noise", e.g. current supply or demand is also affected by random factors, so that all information cannot be inferred from the current price. More generally, as long as the number of relative prices is less than the number of random variables which affect agents' payoffs, REE prices will be partially revealing.

Partially revealing prices create incentives for market formation that could explain the process by which financial intermediaries generate new trades following MS liberalization. If there is noise in current prices so that informed and uninformed agents have different expectations of future prices, there are incentives for trade between the informed and the uninformed. There is therefore a tendency for markets to develop endogenously until the gains to be made from the difference in information between the informed and the uninformed are just sufficient to provide the necessary incentive for costly information collection (see Grossman and Stiglitz [1980]).

If we place financial intermediaries in the role of informed traders, the removal of restrictions on interest rates can be viewed as allowing a return to arbitrage sufficient to provide intermediaries with the incentive to acquire information and trade on that basis with the uninformed isolated agents in MS's framework. Financial intermediaries learn that capital investment will command a high return (or the uses in which that high return is available) and therefore bid on the current market for savings. The higher deposit rate informs all traders that capital held in the form of money will command a higher future return so that they economize on its current use. Price would be bid up until the marginal cost of capital to an intermediary is just equated to the gain from better information.

However, in the MS framework something else is required to complete the story of endogenous development. Although small individual traders learn from price that real money has a high future return, capital indivisibility means that they are unable to amass sufficient capital (in the form of savings) to realize that return. Only financial intermediaries who are "large" relative to other traders will have the ability to utilize the information by pooling the savings of the small units, and such agents are explicitly excluded in REE models.

Even assuming negligibly small agents, REE does not have the prescriptive properties of the AD competitive equilibrium (CE). Under incomplete information Pareto efficiency must take account both of the information which an individual has at the point in time when welfare is measured, and of truth-telling constraints. A particular market arrangement may not be implementable if it depends on private information the individual does not wish to reveal (see Holmstrom and Myerson [1983]). Viewing the REE as an arrangement to extract private information, Laffont [1985] has explored its welfare properties. We restrict our attention to the empirically appealing case of partially revealing REE. Laffont shows that partially revealing REE are not generally Pareto efficient among incentive compatible mechanisms[12] in even the weakest sense of the term. (*Ex post* Pareto efficiency when all information is public knowledge so that no insurance opportunities exist). The reason for this is similar to that described by Grossman: because prices are not fully revealing, the private acquisition of information has a positive externality on the degree of information conveyed by price. Thus the social value of information exceeds the private value and by subsidizing or taxing information acquisition, public intervention may internalize the externality.

REE prices may thus explain how non-regulated interest rates transmit and aggregate information, and it may be just because interest rates only reveal partial information that financial intermediaries have an incentive to collect information and expand trades. However, REE are not constrained Pareto efficient. Further, REE can only provide a very limited guide to the MS process because it is only sensible when agents are small.

---

[12] In these second best situations, one asks whether the equilibrium is Pareto efficient among existing possibilities.

2.4: General equilibrium of an uncertain world with incomplete markets, asymmetric information and money

Preliminaries

We may describe the lagging economies in the MS analyses as subject to significant non-convexities in production, asymmetric information and a shortage of markets in which individuals can shed risk. Neither McKinnon nor Shaw viewed these features as solely the result of government intervention, for example Shaw argued that "in the WV regime with perfect mobility, price flexibility, foresight and competition, a financial system serves no purpose" (p.78) and that intervention was only one reason for segmented markets (p.125). However, intervention prevented the emergence of more viable processes. If interventions are removed banks' information collection and actions can act to inform the price system, allowing prices to transfer and aggregate information so that they signal relative scarcities. Their operations permit trading opportunities to be established and relative prices determined where none had existed before, that is, they help to complete markets. In order to examine the policy implications of this prediction of the liberalization hypothesis, we must move to a general equilibrium context which takes account of market interaction and the influence of expanded markets. The AD economy is the only logical link we have between decentralized economic activity and efficiency. MS may therefore be interpreted as suggesting that the more closely approximated are the complete prices and markets of the AD mechanism, the nearer will an economy move to AD efficiency properties. Indeed, Shaw (p.47) says that "The classic conclusions of the WV (can be called upon) to explain the advantages of the shift in development strategy from repression to liberalization".

In the following we examine the welfare implications of extending the general equilibrium model to incorporate the features described in MS. Extensions have served mainly to incorporate more 'realistic' informational features, and to allow for incomplete markets (prices are not determined for every finely defined commodity), so that there is a rationale for active markets at every date. In these applications the REE has replaced the AD equilibrium concept. Money is considered next. The prices MS focussed on were rates of return on money holding, but the neoclassical model which is the rationale for liberalization not only does not include money but makes

money redundant. It is not a coincidence that allowance for observed institutional features provides a role for money in the general equilibrium system. This is almost the insight of MS: when economies do not conform to the idealization of the AD world, money becomes important. The irony is that one at the same time loses the clear prescriptive insights of AD.

Incomplete markets with asymmetric information

The AD model deals with uncertainty by making very fine distinctions among "commodities": the same physical good is a different commodity according to its location and date of availability or use, and the state of the world in which it is made available or used (the well-known state-contingent commodity). The price ratios between each of these finely defined commodities are determined in the single 'marketplace' which takes place at the beginning of the economic system. Agents, taking these prices as given, are therefore not subject to uncertainty about the present value of producer plans or about consumer budget constraints. Producers are unaffected by uncertainty or risk because all possible inputs and outputs of the production system have a firm price. Given their probability distributions over possible states of nature and their risk attitudes, consumers' buying and selling of the dated state-contingent commodities serve an insurance function.[13]

This seems the key to one idea behind Shaw's view of financial intermediaries. Complete markets remove risk by providing prices over events/dates; liberalized financial intermediaries also remove part of the risk borne by producers by making price judgements on the basis of better information (Shaw [1973], p.127). However, the MS environment and information structure cannot be accommodated in the AD model. The AD model does not allow for the differential information[14] among agents which MS stressed as a crucial component of the risk faced by agents in LDCs. Radner (see [1982a]) has extended the AD model to allow for differential information among agents. If the feasible plans of the agents with incomplete information are restricted to those in the full information set (i.e. in the set of plans available if an agent had full information)

---

[13] Complete asset markets at the initial period and spot markets thereafter serve the same purpose.

[14] In fact, the AD equilibrium requires that at the opening market agents have common forecasts of equilibrium spot prices in the future for every event.

which are compatible with the given information, existence and efficiency of the extended model, relative to the given information structure, can be shown. In addition to the usual convexity and continuity assumptions, three conditions must hold for this result (Radner [1982b]): the information available to an agent must be independent of his or any other agent's actions, there must be no moral hazard, and agents must not use equilibrium prices to make inferences about other agents' behaviour. Price-taking agents, complete markets, and a single pre-history determination of prices continue to be assumed. Thus, while allowing for asymmetric information, most of the features associated with its existence are excluded in order for the extended AD model (Radner's [1982a] terminology) to give AD-type results. It is also evident that the extension can be of no use in justifying the MS arguments because its existence and efficiency proofs depend on the absence of those features which they stressed; namely, bank acquisition and use of information with resource expenditure, and the use of prices to signal information about capital scarcity.

The single opening of markets in the AD framework has its explanation in the presence of complete markets and no transactions costs since there is no advantage to be gained by re-opening markets. The institutional features in MS have no place in the framework. If all prices are determined and accounts are settled at the beginning of time, agents have no need to economize on "search and bargain" by holding money, nor would money be held as a store of value. Consumers who know their net present values need not hold shares so no stock market would exist. In addition, while the extended AD provided a means of dealing with asymmetric information it could do so only by restricting their feasible plans.

In order to incorporate these real financial features and provide a more satisfactory characterisation of asymmetric information[15], general equilibrium theory has adopted an approach which looks for REE in a sequence of incomplete markets. With agents using equilibrium prices to make inferences about the environment, an REE[16] is a set of current prices, common price ·

---

[15] The recognition of asymmetric information is sufficient to necessitate a sequence of markets (Arrow and Hahn [1971]). When commodity availability or use depends on the state of nature and a market participant is aware that some traders have information which he lacks, he will be unwilling to enter into conditional contracts. Future prices which depend on the information will be treated like a random variable. As a result it becomes appropriate to consider markets as opening in sequence as information becomes available. Similarly, the individual may invest in securities to insure against future uncertainties.

[16] Which originated in Radner [1972] and Lucas [1972]

expectations and consistent plans such that, given current prices and price expectations, each agent's plan is optimal for him, given his sequence of budget constraints. The sequence economies with incomplete markets is about as far as formal neoclassical theory has gone in allowing for the type of environment and conditions in which liberalization is to be effected. The incomplete nature of the theory is indicated by the difficulties encountered in proving RE equilibrium[17]. Allen ([1986], p.20) stresses that "the existence problem for rational expectations equilibrium is definitely still not fully solved in a satisfactory manner". Allen summarizes and discusses some of the existence (and non-) results. It is sufficient here to give the flavour of the latter by citing the result from Hart's [1975] exchange economy example of nonexistence: intuitively, equilibrium failed to exist because incomplete markets produce noncontinuity in budget constraints and hence demand functions (correspondences) so that the usual fixed point techniques could not be used.

Existence problems imply the tentative nature of results but the welfare analyses of REE in incomplete markets produce two points relevant to a prescriptive assessment of a liberalized system. Hart [1975] showed firstly that REE are not generally optimal in economies where the market structure is incomplete and secondly that, unless all other markets are available, the opening of an additional market need not produce a Pareto improvement. For the first demonstration it is not general but constrained Pareto efficiency (PEy) that is sought. One asks whether the REE allocation reached is PE relative to allocations that can be achieved through the existing set of incomplete markets. Hart gave an example where a Pareto-dominated equilibrium may be the only one attainable [18] because the move to a Pareto superior allocation would require a change in prices and expectations which could only occur through trade. However, that trade is not possible because markets are incomplete (in Hart's three-period model there is no borrowing and lending

---

[17] Geanakoplos and Polemarchakis [1986], and others, have obtained existence in 'financial securities competitive equilibria'.

[18] This result has since been rigorously extended. Geanakoplos and Polemarchakis showed that, generically, equilibria of pure exchange economies with incomplete markets are not constrained PE. Later work (with others) confirms this result when production is incorporated. Quinzii [1988] provides an intuitive explanation. Imprecisely, the planner has an advantage over the market because he realizes that a relative price change will affect welfare since marginal rates of substitution are not equated in incomplete markets. And by marginal changes in allocations and production decisions, the planner can change relative prices.

and no futures market). The equilibrium reached then depends on the prices at which the economy starts: there is insufficient trade to provide the 'market forces' which would lead the economy to a PE allocation in the existing set of markets. This could be the MS argument: the equilibrium achieved in the repressed economy is inefficient because the repression of prices does not permit market forces to reach the growth-promoting allocation. MS therefore postulate that expanding borrowing and lending possibilities through financial intermediation can improve efficiency.

The second point made by Hart disputes (refutes?) this conclusion by considering the introduction of a new security in the incomplete market structure at the first date and finding that its introduction in fact makes everyone worse off. If utility from consumption is not constant across dates, the time at which trade takes place becomes important (recall that Hart is analyzing an exchange economy). In Hart's example the additional market implies that price-taking consumers will trade at its opening until all gains from trade are eliminated, without recognizing the interdependence of gains from trade at different periods. Utilities from consumption in later periods are sufficiently reduced relative to the situation without the additional market to make all consumers worse off.

Although it may be stretching credibility to attempt to apply this very specific abstract result to an empirical situation, we can attempt to relate it to an argument put forward by Caribbean economists when urged to raise deposit rates in order to provide banks with liquidity for the finance of real investment. The argument says that "there is no necessary correspondence between real and financial saving" (Worrell [1984], p.60) - an increase in financial saving may be translated into loans for consumption rather than investment purposes. In terms of the theory and MS's framework, one can recast the argument as follows. In a repressed regime encouragement of the deposit market is equivalent to adding a market - providing consumers with an additional security (deposits) in which they can trade. But individual consumers do not perceive their future consumption possibilities from general investment (the connection between earlier and later trades), and given the opportunity to trade they will exhaust the gains from trade at later periods: current financial saving is borrowed for later consumption (say at date 2), rather than invested in

real assets at date 2 for consumption of the return at date 3. And banks are concerned only with their individual profits, not with lending to improve investment. As Hart points out, an economy with incomplete markets is in a second best position. Only if all markets are opened can an overall improvement be expected. (In our incredible example venture capital opportunities may be a possibility).

As emphasized before, these sequence economy models are significantly different from the MS characterization in their assumption that agents' information structures are exogenously fixed. This is an unattractive assumption from the 'reality' viewpoint but it has not been relaxed because in general the technology of the acquisition and use of information does not satisfy the assumptions required for theorems proving existence and efficiency. Both McKinnon and Shaw stressed that acquiring and using information involved the expenditure of resources. Similarly, Radner [1982a] notes that a production plan which requires more information must include increased inputs. Radner and Stiglitz [1984] show that there is a fixed set-up cost to information acquisition which introduces a non-convexity into the production possibility set. As usual this implies that there is a discontinuity in the demand for information, hence the difficulty with standard existence and optimality proofs. Radner and Stiglitz note that in areas where information is important specialization may therefore be common. Radner ([1982a], p.974) also points out that a producer may have different information structures available, each with a production set, his total production set being the union of the different sets, and that union may not be convex. These results pose a further problem for a neoclassical rationale for liberalization. As already pointed out, MS posit an economy with nonconvex production possibilities but appear to believe that these can be mitigated by information-gathering banks. However, the Radner-Stiglitz theorem implies that the solution is itself likely to introduce further nonconvexities.

Introducing Money

So far we have considered the implications of introducing incomplete differential information and incomplete markets into the standard neoclassical GE model. These are both features of the MS environment. In MS freely operating financial markets are seen as the corrective for

resulting imperfections. Here we look at the explicit consideration of money and ask whether a more rigorous treatment of GE in a monetary economy supports their intuition. We conclude that it does in important positive respects, though not from a policy perspective.

Two features of the AD model preclude a role for money (see Ostroy [1989]): the completeness of the markets and the single budgetary constraint of its agents. Since markets are complete, the equilibrium which coordinates the actions of all agents is completed at the first date and markets need never re-open, although actual delivery and production may take place sequentially over time. There is obviously no need to maintain a store of value or medium of exchange (or any other financial assets/institutions) since complete (and implicitly binding) agreement has been reached on the precise commodities to be exchanged over all time and in all eventualities. Suppose markets were not complete, but re-opened every period for trade (i.e. there are a sequence of markets). As long as agents satisfy only a single multiperiod budget constraint so that trades in any period need not balance, an asset with no intrinsic value would still have no place. Agents would be concerned only with equalising their overall inputs and outputs and, with rational expectations, they know that these will be consistent. With a sequence of single period constraints, however, the need to meet a budget constraint each period would frustrate some trades so adding money would be like completing markets.

Gale [1982] views money's role as based on even more fundamental features of the AD market. An AD equilibrium can be reached at date zero only because when trades are made agents have sufficient trust in each other to be confident that the arranged deliveries will take place (or there is some implicit enforcer operating). In reality of course, individuals are not that trustworthy. Specifically, in the final period of the market game, the individual who has previously received, and now has to make a delivery, has an incentive to retain the agreed commodity. A sequence of budget constraints can act to ensure sequential delivery since it requires balanced trades every period: there is, so to speak, a check on agents' contributions. But this requires that agents borrow and lend between periods and have some means of transferring wealth between markets. Certain trade patterns could not be accomplished if budgets had to be balanced at every date. Bonds could not serve the purpose of storing wealth because they require trust in the issuer.

Money would facilitate trade because its use could act as a guarantor of forward delivery if, as in Gale's example (op. cit., p.239), agents were issued with fiat money which had to be returned to the issuer at the end of the economy, but could be freely exchanged among agents in the interim. The agent making the forward delivery would have to fulfil his contract in order to retrieve the money he had previously paid for commodities received. However, money can only serve this purpose because the arrangement is enforceable by the fiat money issuer. And if this power of enforcement exists, it could also be used to enforce commodity deliveries. But in the latter case the number of contracts the authority would have to monitor and enforce would be very great. When treated as if it had a uniform value and used to balance budgets, money acts as a store of information because it allows the issuer to keep track of what agents have obtained from the economy. The use of money is therefore less costly than the monitoring required in its absence: money has the advantage of being a decentralized method of facilitating trade. As such, Gale characterises the monetary arrangement (which includes both the positive value attached to a paper asset of no intrinsic value and the enforcement power) as a social institution which acts to complete markets, that is, permit more trades.

Gale shows that when there is a complete set of paper assets (one for each state of nature) the equilibrium of the monetary economy is in the sequential core. But if the set of paper assets is incomplete (and this requirement is as unattainable as that of complete markets for securities), the equilibrium allocation is almost always Pareto inefficient and as a result the sequential core is normally empty. With incomplete assets agents have different marginal rates of substitution between consumption in different states, giving coalitions an incentive to form to change the allocation.

Incomplete markets are constrained Pareto inefficient. Money helps complete markets but unless there is complete money the equilibrium is not PE: we already knew that adding markets does not necessarily make everyone better off. We go on to consider whether financial intermediaries help.

2.5: Financial Intermediation Mechanisms

Gale's account of the role of money dealt only with fiat money: government is required to enforce the tax payments which act as a decentralized guarantor of future delivery. However, outside money is dominated by inside money in most monetary systems, and the financial liberalization argument is primarily concerned with the benefits to be gained from intermediation. It is the actions of financial intermediaries, rather than money *per se*, which act to convey information. In this section we describe some of the literature which derives financial intermediaries and their contracts as endogenous market responses to private information in markets where producers must obtain outside financing. In all of these, as in MS, the *raison d'etre* of the financial intermediary lies in some form of scale economies which serve to make it the least cost means of overcoming the agency problem resulting from differential information.

These models provide formal justification of MS's insights but they also show that, while financial intermediation is the least cost means of providing additional markets in a situation with asymmetric information, it results in agency conflicts of its own whose resolution may require an exogenous mechanism. That is, the equilibrium with the financial intermediary mechanism may be constrained Pareto inefficient. However, the behaviour described by these models also make it clear that the conventions, contracts and processes observed in financial intermediation can be viewed as facilitating exchange and production in an incentive-consistent fashion. Ill-conceived official intervention may interfere with these mechanisms.

The earliest of these information-based treatments of financial institutions, Leland and Pyle [1977], recognize, like MS, that financial markets are particularly afflicted by problems of asymmetric information and are unlikely to allocate finance efficiently unless information is reliably conveyed. Unlike MS, they also recognized that asymmetric information about project returns could produce adverse selection problems. Information flows are improved if specialized financial intermediaries emerge to exploit economies of scale in the collection and sale of information. Leland and Pyle did not settle the next question regarding the quality of the financial intermediary's information. Further, the question of how financial intermediaries obtain informa-

tion is not addressed.

Diamond [1984] shows how incentive compatible contracts can answer both these requirements. His analysis proceeds by addressing the levels at which asymmetric information becomes a problem. When the realized outcome of a project with a random return is private information to a borrower the standard debt contract (fixed repayment, bankruptcy penalty when insolvent, with lender keeping the residual) is the optimal arrangement between a borrower and lender. A fixed repayment (independent of the realized state) removes the incentive to misrepresent a favourable state of nature. The bankruptcy penalty and principal-take-all clause ensures that the borrower will only declare insolvency if this is in fact the case (i.e. there is a deterrent in the shape of a cost of declaring insolvency). However, this optimal contract is costly because it entails a positive probability of bankruptcy. This could be avoided by monitoring project realization, but the privacy of project realization implies that each security holder (lender) would have to monitor, resulting in effort duplication or a free-rider problem. The solution is to delegate the monitoring to a single agent: this in turn generates an agency problem which can be solved by a standard debt contract between the delegators and the monitor (depositors and bank). The financial intermediary mechanism works here because its net costs are lower. Although provision of the correct incentives for delegated monitoring is still accomplished through a costly bankruptcy threat (a nonpecuniary cost borne by the borrower with positive probability), if the intermediary contracts with many firms with independent, identically distributed (i.i.d) risks, the probability of bearing the bankruptcy costs because of a single firm's failure tends to zero. Average delegated monitoring costs decline with numbers monitored so that financial intermediation works because these economies of scale compensate for the costs of incentive provision.

Diamond (and Leland and Pyle) thus give operational content to the MS view of the informational role of financial intermediaries. The intermediary is the least cost method of information production because with diversified assets it is able to collect and monitor information in an incentive compatible (reliable) manner. Leland and Pyle assumed economies of scale in information collection: some such assumption must be made to explain why it is advantageous for agents to trade via intermediaries rather than directly. But if the viability of financial institutions

depends on their size, we should expect them to behave strategically.

Gale and Hellwig [1985] [19] also demonstrate that the standard debt contract is the optimal contract between borrower and lender when there is *ex post* asymmetric information (project outcome is private information of the borrower). Intermediary observation of project return which occurs when repayment is not possible is interpreted as bankruptcy which imposes a nonpecuniary cost on the borrower. Avoidance of this cost produces credit rationing: the optimal contract keeps investment below its first best level (that under perfect information) because with diminishing returns to investment, lower investment increases profits and reduces the probability of bankruptcy. Again we have a situation of asymmetric information corrected by a contract whose incentive compatibility requirements prevent a first best outcome.

Several models based on Diamond and Dybvig [1983] have considered the demand deposit contract separately, concentrating on the financial intermediary as asset transformer and hence insurer rather than information collector and monitor. By providing liquidity, banks are able to insure agents who learn their preferred consumption profile privately. But, because the deposit contract provides liquidity by enabling banks to supply deposits at a pace that is out of step with production, there is an equilibrium where all depositors try to withdraw early, resulting in a panic run on the bank. Either a suspension of convertibility or deposit insurance could stifle the panic, by assuring would-be withdrawers that funds will be available. Later models using the Diamond and Dybvig framework[20] derive runs from basic bank characteristics such as depositors' limited information about banks' assets. Bank runs are undesirable because the intertemporal allocation of investment resources is suboptimal (production is interrupted) when depositors withdraw early. We see again that the bank arrangement for services (here insurance) may be improved upon by a planner.

One analysis without this result is that of Williamson [1988] where bank 'failures' are only associated with particular states of nature, the allocation remains optimal and there is no role for

---

[19] See also Williamson [1986]
[20] See, for example Postlewaite and Vives [1987], Jacklin and Bhattacharya [1988]

government intervention. Williamson's failures [21] do not have the 'flavour' we would normally associate with bankruptcy in financial institutions: rather than being an otherwise undesirable response to adverse circumstances, they occur by *ex ante* agreement among members of banking coalitions in states of the world in which capital market trades are preferable to complex banking structures. It may be more fruitful to view Williamson's bank and no-bank equilibria as delineating the conditions (given his model) under which banking coalitions are optimal mechanisms (when assets are illiquid because of asymmetric information).

The role of deposit insurance as a means of avoiding bank runs, as well as the incentive problems it creates, are well-recognized in the liberalization literature (see Shaw [1973], pp.64-66 and The World Bank [1989], p.76). The models sketched above provide a formal description of quite widespread notions. They also show that in averting runs deposit insurance may actually be Pareto improving . It is because private incentives on the market create externalities (in the Diamond and Dybvig model early withdrawers do not take account of the social costs of destroyed risk sharing and interrupted production) that interventions such as deposit insurance may have a role. This externality feature is not recognized outside of the formal literature. The 1989 *World Development Report* suggests that

> 'by providing implicit and explicit deposit guarantees, governments ... have suppressed the market forces that otherwise would have eliminated or reorganized unprofitable firms and allocated the associated losses.' (p.79)

In this view runs are appropriate discipline rather than the costly result of random disturbances, and their possible costs are not set against the expected costs of insurance.

As more 'realistic' assumptions about the distribution of information are added on to the Diamond-Dybvig model, reserve requirements also acquire an explanation in terms of the internalizing of an externality. Simplifying their analysis, Bhattacharya and Gale [1987] showed that, although banks may have an incentive to create an interbank market in which they may share liquidity risks, if their investment and the proportion of early withdrawers is not publicly observ-

---

[21] Williamson addresses this issue himself, though not entirely satisfactorily, through appeal to stylized facts of bank failure.

able, individual banks will reduce their investment in liquid assets and rely on the interbank market. Hidden action creates a free rider problem. A legal reserve requirement on which the full rate of interest is not paid, together with a discount window at a subsidized rate of interest, solves the second-best risk sharing problem: the financial intermediaries who learn that they are subject to a larger proportion of withdrawals can take advantage of the arrangement.

It should be pointed out that there are more general models concerned with the endogenous derivation of financial intermediaries which find less of a role for government intervention. Thus Boyd and Prescott [1986] show that financial intermediaries are a constrained Pareto efficient mechanism in an economy where investment opportunities are *ex ante* private information, a signal of which can be acquired at a cost. Individual endowments are insufficient to both undertake investment and evaluate investment projects. Agents with 'bad' and 'good' projects self-select for coalition membership or project evaluation and the resulting core equilibrium is supported by coalitions which must be large in order to ensure available financing for the good projects. Again we have the size requirement for intermediaries - here the possibility of monopoly power is avoided by the formal framework: the economy and each intermediary have a countable infinity of member agents.

2.6: Conclusion

We have approached this discussion as if liberalization would start from a clean slate, that is, as if considering an economy with incomplete markets, uncertainty and asymmetric information in which money and financial intermediaries are incorporated at the start of history and where there is no government intervention (except in so far as this is implied by the introduction of money). In reality, liberalization would usually follow a long history of regulation: institutions and expectations will have formed. In practice, therefore, all the results above would need to be reinterpreted in the context of the particular economy and era in which reforms are introduced. Neoclassical theory indicates only that we cannot expect constrained Pareto efficiency in a liberalized regime, given the environment which it has been argued necessitates liberalization, It is conceivable that unadulterated deregulation in a given economy would produce a situation

which could be described as a Pareto improvement.

However, many of the imperfections MS attributed to repression could just as readily be explained by the environments of asymmetric information, indivisibilities etc. as we show in the following chapters. Price dispersion is consistent with an unregulated market where information is costly since the cost of acquiring information reduces the net benefit of searching for the lowest price (see Stiglitz [1977] and Salop and Stiglitz [1982]). Market power can result from imperfect information and search costs (see Chapter 4); credit rationing may be explained by decentralized banks dealing with asymmetric information (see Chapter 7); entry may be frustrated or impeded by imperfect information (see Chapter 6). This makes policy prescription difficult.

Unfortunately the focus on high interest rates has masked most of the more subtle arguments underlying the MS recommendations. It omits the crucial informational issue and leaves policy makers with the not unjustified perception that high rates of interest serve only to give free rein in exploiting oligopolistic advantages. As one Caribbean central bank governor has argued (Blackman [1985], p.35):

> 'In developed countries the market mechanism can be relied upon to allocate resources in a satisfactory manner in the vast majority of cases. In developing countries the existence of monopoly, oligopoly or other market rigidities will frequently cause the market to operate in a perverse manner. It does not help to behave as if markets are efficient when in fact they are not. For example, there are only eight commercial banks in my country, seven of which are foreign banks and hold 81% of banking deposits. They cannot be left to fix interest rates among themselves.'

In its stress on prudential regulation of the banking system, the 1989 *World Development Report* recognizes the need to supplement deregulation because of the asymmetric information between depositors and banks, in the case where the bank is agent and the depositor principal. But, given the other sources of asymmetric information in the financial sector, we show that the tax/subsidy measures a central planner could usefully effect go beyond purely prudential controls on financial intermediaries.

Chapter 3: Financial Markets in the Caribbean Commonwealth: Four Examples

3.1: Introduction

This chapter describes key features of the financial markets and regulatory environment of the four Caribbean Commonwealth countries which provide the empirical framework of our analysis[1]: The Bahamas, Barbados, Belize and Trinidad and Tobago[2].

Observations in Caribbean financial markets suggest several distortions. To take those on which we will be focussing: interest rates are low and inflexible compared to rates on alternative foreign markets, although the economies are small and open, and most have no explicit interest rate controls. Potential borrowers, especially those of small scale in non-traditional activities, complain of loan rationing. There have been several unsuccessful attempts at entry by financial firms. Conventional analyses of LDC financial markets in the McKinnon-Shaw tradition suggest that these outcomes result from regulation. Caribbean analysts offer explanations based on bank behaviour. They view low and inflexible rates as the result of monopoly power. Intermediary credit allocation is seen as reflecting bank familiarity with particular sectors, rather than rates of return. The failure of potential competition is seen as symptomatic of the entrenched position of long-established banks. The models in the following chapters will examine the outcomes expected from liberalizing financial markets with typical LDC characteristics - poor information systems and small numbers in the modern sector[3] - using these hypotheses, suitably interpreted, as our starting point.

The financial systems we will describe reflect the effects of interest rate regulations, credit policies etc. The liberalization process would involve removal of these measures. We will therefore be conducting a counterfactual conceptual experiment. In order to do so we specify two sets of 'stylized facts'. The first set describes the 'institutional' framework; the second the

---

[1] This choice is dictated by the availability of descriptive material . Other countries of the Commonwealth Caribbean are expected to have very similar systems. The information available even for these four countries is not uniform and therefore the bases used for comparison will only be approximately comparable.

[2] Trinidad and Tobago is one state of two islands.

[3] Of course, small numbers are an inherent feature of Caribbean economies.

behavioural observations we wish to account for. These are described in Sections 3.3 and 3.4, respectively. Section 3.2 outlines relevant characteristics of the four economies. Summary descriptive data on the economies and their financial sectors is given in Tables 3.1 and 3.2.

3.2: Notes on the individual economies and their financial sectors

The four economies are all highly dependent on foreign transactions, principally in the form of invisibles, although petroleum has been an exception in Trinidad and Tobago. The banking habit is widespread, with all four having a money to GDP ratio of about 40% (see Table 3.2). The unusual feature common to their financial sectors is the predominance of multinational banks. Specific features of each economy are described below.

The Bahamas

The financial sector in the Bahamas differs from those in other islands because of the country's status as a tax haven and major offshore financial centre. Financial institutions locate there to take advantage of tax measures and administrative arrangements, as well as to service clients benefiting from these. Finance and tourism are the most important productive sectors. However, limits on convertibility and exchange controls continue to insulate the domestic portion of the banking sector from the external market. Our concern is with domestic operations. Financial organizations with both resident [4] and authorised dealer status [5] carry out Bahamian dollar business and foreign exchange dealings (the latter may include offshore business) but tend to view their domestic and external operations separately (Ramsaran [1984]). The two operations would be connected by foreign currency loans made to residents and by external transactions between local institutions.

In 1978 there were ten resident commercial banks; six of these being branches [6] of

---

[4] Resident banks are those with a licence to carry on Bahamian dollar business i.e, they are resident for balance of payments and exchange control purposes.

[5] Authorised dealer status is normally accorded to banks when countries operate exchange controls. It allows the banks to hold gold and foreign exchange and to carry on the foreign exchange transactions necessary for the usual trade and travel purposes. Capital and financial transactions would normally require explicit exchange control permission.

[6] Branch banks have no legal status separate from that of the multinational parent bank, they have no assets and liabilities of their own, though separate accounts are maintained for internal accounting purposes. A

multinationals, three subsidiaries, and one a consortium subsidiary. However, several of these banks are essentially concerned with external business rather than domestic intermediation: at the end of 1978, 96% of Bahamian dollar loans and 98% of deposits originated in the seven clearing banks. With 52 offices maintained by the clearing banks, population per bank office was 4,500 in 1979. Banks are subject to a cash reserve requirement and liquid assets ratio on Bahamian dollar deposits. In 1978 no liquid asset ratio was in effect but holdings of assets which would have qualified amounted to almost 30% of Bahamian dollar deposits (Ramsaran [op.cit.], p.87). Other financial institutions include insurance companies, finance companies (which take deposits to on-lend for mortgage or consumer installment credit), trust companies (which principally provide trustee, company management services etc.), the Post Office Savings Bank and the People's Penny Savings Bank. These last two are negligible in size (in 1978) compared to the others.

Barbados

Barbados is also an offshore financial centre but the enabling legislation is very recent and it lacks the tax haven advantage of the Bahamas. Consequently the activities of resident and non-resident branches of the same multinational are clearly demarcated: resident branches do not conduct offshore business. Major areas of economic activity are tourism and manufacturing, with the sugar industry remaining a relatively important source of net foreign exchange. There are seven commercial banks. Five are branches of multinational banks, one is government-owned, and the other is locally-incorporated and owned by a regional insurance company. Commercial banks are required to hold a portion of their deposits in cash as well as in stipulated government securities. In 1986 the cash reserve requirement was 8%, and stipulated government securities 22%, of deposit liabilities. Although this requirement is substantial, the banking system as a whole has held very substantial excess reserves ever since the requirement came into effect. During 1985, for example, excess securities averaged almost 30% of requirements. Selective credit controls have been in effect since 1977. These place limits on commercial bank credit to the distribution and personal sectors, with certain housing-related credit exempted. Consumer installment credit is

---

subsidiary is a legal entity distinct from the parent bank.

also restricted with exemptions applying to commercially-related purchases e.g. for taxis. Both minimum (for deposits) and maximum interest rates have in the past been fixed by the authorities.

Other financial institutions include insurance companies, trust companies (which were established by the commercial banks to handle their mortgage finance and trust business), finance companies (which take deposits to fund consumer instalment credit) and credit unions. In 1983 deposits at the latter represented only about one percent of deposits at commercial banks. There is also a government-owned mortgage finance company.

Belize

Belize's economy is agriculture-based; sugar, bananas and citrus fruit are the main crops. Almost 50% of the population live outside the urbanised districts. There are four commercial banks, three of which are branches of multinational banks while the fourth, although 50% locally owned, is an affiliate of an American multinational and has its head office located in Honduras. Banks are subject to a 20% (of deposit liabilities) liquid asset ratio, and a 5% reserve requirement. Selective credit controls have been used to reduce credit to the distribution sector. There were 16 bank offices in 1983, with only one of the 6 districts not having a commercial bank office; the average number of customers per office was 9,085. In addition to insurance companies and a government savings bank, the financial sector includes cooperatives and credit unions. Among the countries considered, Belize is the only one that appears to have an unorganized money market worth mentioning (Luben [1983]).

Trinidad and Tobago

The economy of Trinidad and Tobago was transformed by the oil boom of the 1970s. The resulting increase in income and wealth has contributed to an expansion of financial institutions. More recently, decreased petroleum earnings may have contributed to the problems encountered by some of these institutions. Finance companies, whose portfolios grew substantially during the boom period, concentrated on credit for the purchase of particular commodity categories (Bourne [1982]) and declining incomes adversely affect both their deposit inflow and the viability of such loans. There are seven commercial banks, two of which are locally owned (one is state-owned).

The others are subsidiaries of multinationals as a result of a government requirement, since 1970, that foreign-owned banks have 51% percent local equity participation. At the end of 1982, banks were subject to an effective cash reserve requirement of 16.1%; they must also hold secondary reserve cash assets (5% of deposits in 1982) at the central bank. Selective credit guidelines have also been used to restrict personal sector credit. In 1982 there were 112 bank offices, with average population of 10,077 per bank office. In addition to the finance houses, there are insurance companies, building societies, and credit unions. Trinidad and Tobago is the only Commonwealth Caribbean country, apart from Jamaica, which has a stock exchange. In 1981 there were eight registered stockbrokers, but no secondary market for corporate securities (Bourne, op.cit.).

3.3: Stylized facts about institutional structure

In the small very open economies of the Caribbean, where domestic financial institutions are predominantly foreign-owned, and export-oriented service industries like tourism and international finance are the major sources of value added, one would expect domestic banks to be both price- and quality-takers. On the other hand, our partial equilibrium models are formulated on the premise that domestic banks fix their prices, quantities, contracts and service independently. This apparent anomaly is explained by the regulatory and institutional structure of the economies; a structure which we describe by two stylized facts:

1. Domestic financial sectors are insulated from international financial markets by exchange controls and nonconvertible foreign currencies;

2. Whatever the ownership structure of the domestic banks, price, quantity and quality decisions are taken by locally-based bank management to maximize domestic profits.

1. Insulation from foreign financial markets

Exchange control regulations tend to insulate the financial sector from foreign markets. In general, financial institutions cannot lend to non-residents and do not compete (legally) with foreign banks for domestic deposits. Residents can only borrow or invest abroad with exchange

control permission which is given only after consideration of each specific case. The non-convertibility of regional currencies obviously reinforces the insulating effect, although, in at least the Bahamas and Barbados, the relatively widespread availability of US dollars, through the tourist industry, does mean that residents have access to foreign currency outside legal channels. In the Bahamas, US currency circulates with the domestic currency. However, it has been estimated that the aggregate value of circulating US dollars represents less than 5% of the value of local currency at any time (Ramsaran [op. cit.] p.115). In Barbados, the proportion is probably negligible. The question of interest in this regard is whether savers purchase US dollars domestically to invest abroad, and do so in sufficient quantities to influence banks' behaviour. In Barbados, the transactions costs and legal risks of accumulating sufficient foreign currency would suggest that this can be answered in the negative. In the Bahamas although an affirmative answer is more likely, two factors should be taken into account. First, residents (mainly foreign exchange earners) are allowed to hold foreign currency deposits, under certain conditions, for current transactions. At the end of 1978, such deposits were 8.2% of total Bahamian dollar deposits. Second, these depositors, given the structure of the economy, are probably those most likely to consider foreign investment but, since they already have foreign deposits, they would have less incentive to attempt black-market-type dealings. About half of such deposits are held in interest-bearing accounts, and not only would this permit the depositors to obtain returns comparable to those on external markets, it would also permit the banks to segment the deposit market, avoiding competitive pressure on local deposits. Support for such segmentation is found in the deposit rate: in 1978, the average fixed deposit rate in the Bahamas moved in the range 5.06-6.79%, while US 6-month CD rates moved between 7.23-10.76%. Ramsaran [1984] also observes that authorised dealers in the Bahamas have often ignored the restrictions imposed on their use of foreign currency to finance local operations, but competition in this domain did not appear sufficient to induce loan rates adjustment in line with foreign rates: the Bahamas prime rate remained at 9.5% between 1974 and 1978, in a period when the US prime rate moved successively, as follows: 7.26%, 6.35%, 7.75%, 11.55% (end of year rates for 1975 to 1978). An even more striking lack of response of local rates to international rates is evident in 1980 to 1982: the

average deposit rate in the Bahamas never rose above 7.58% while the average US 6-month CD rate was over 12% for most of the period. The prime rate, which banks would be expected to raise, remained at 11% for the entire period while US prime rates reached record levels - over 16% throughout 1981. The experience of the Bahamas in this respect is interesting because there are no explicit interest rate regulations there, and the inertia of their interest rates is what would be expected in a collusive arrangement.

Misaligned exchange rates provide an incentive for savers to hold funds abroad. For small misalignments, the net benefits of holding foreign funds (reduction of the potential loss from domestic devaluation less the transactions cost of illegally purchasing and repatriating foreign funds) are probably not sufficiently great to destroy the markets' segmentation. Large misalignments inhibit banks' ability to compete with foreign markets since attempting to do so would require interest rates high enough to compensate depositors for the insecurity of holding domestic funds. However, there is little evidence of misalignment in the Caribbean, except for a brief period in Trinidad[7]. Even illegal trade in foreign currency takes place at the official rate.

2. Local Decision-making

With the exception of Trinidad and Tobago and Jamaica, where 51% local ownership of commercial banks is a legal requirement, banks in the Commonwealth Caribbean are mainly wholly-owned branches of multinational commercial banks. Most have regional offices located in the region (for example, in Bridgetown, Nassau, or Miami) which cover several islands. General policy (such as sectoral emphases in the loan portfolio) is set by regional head office which also has to approve very large loans, but other decisions are taken locally, within guidelines set by the head office. Specifically, price and quality decisions are taken by the senior local bank manager.

In Trinidad (and Jamaica) it can be assumed that the subsidiaries continue to make strategic decisions on a basis similar to that of the branch banks, since the management remains part of the

---

[7] Guyana and Jamaica are the notable exceptions to this claim. The economic disarray in these countries during the seventies and eighties would have so altered agents' perception of structure that very specific models would be required to investigate liberalization's effects.

multinational personnel structure. In Trinidad and Tobago, local equity is very diversified so that owners are unlikely to marshall sufficient votes to exercise control over decisions. While the National Insurance Board holds 10% of equity, it is itself controlled by a triumvirate of government, trade unions and private business. Thus the subsidiaries' strategy variables and decisions are unlikely to differ significantly from those of multinational branches. Of Jamaica, Baum ([1974] p.15) said:

> "The result of Jamaicanization has been only a passing of theoretical equity participation. The bank as an international institution remains in control; the management exercises control and is paid for the job it does. The shareholders have an interest in the profits but are not in an effective position to substitute their collective judgement for that of the bank's management."

Furthermore, even the government-owned commercial banks in the region continue to behave very similarly to the multinational branches.

We then have a decision structure where locally-based management is taking decisions on the basis of 'guidelines' set by multinational head offices. Or the management of domestic banks are taking decisions on the basis of guidelines set by their local board of directors. In either case we assume that the goal is profit maximization. What would be the force of the guidelines? Historically, the guideline set for interest rates was the London Bank rate (McClean [1975]), at a time when a currency board system operated and local currency was tied to sterling. Not surprisingly, in view of the date of British Competition and Credit Control, this appears to have come to an end in 1971. Subsequently, with increasing U.S. and regional influences (a Caribbean common market was formed, for example), profit maximization required that greater attention be paid to domestic conditions, and head offices now rely on the judgement of locally-based management, so guidelines appear to leave wide discretion to such managers. We can therefore assume the strategic interaction of the banks with regard to all decisions (with the exception of entry) to be local interaction within the environment of the individual economies.

Banks have been subject to a considerable degree of regulatory control since the advent of central banks or monetary authorities. In addition to the standard reserve requirements, selective credit controls limit their choice over loan portfolios. But, with the exception of Barbados, the

authorities have in general allowed the banks to determine their own interest rates. In Barbados banks' interest rates are largely determined by usury laws and central bank regulations, or are subject to central bank approval. However, the size and resources of the banks relative to the local economies, and their crucial role in those economies, does mean that central banks have to take account of the commercial banks' interests (a point made by Bourne [1984]). Commercial banks continue to represent a source of foreign exchange (foreign loans to local companies would normally be negotiated via local branches) and represent financial expertise not available locally. Although managers may be nationals, they tend to be experienced employees of the multinational bank and many would presumably leave with the exit of the multinational. Generally speaking, there is no domestic financial intermediary which could substitute . Thus, central bank regulation is not likely to constrain bank behaviour beyond some point the banks themselves find acceptable. This point remains relevant to the situation without controls.

3.4: Stylized facts about behaviour

The introduction indicated that Caribbean analysts have deduced certain characteristics and expectations from their observations of and experience in these markets. Our analyses demonstrate that the behaviour they inferred is consistent with strategic behaviour under imperfect information. These inferences may be summarized by the following:

3. Commercial banks dominate and are the price leaders in the financial sector, with few substitutes for either savers or borrowers.

4. Banks' allocation and contract decisions are strongly influenced by the knowledge of certain customers and sectors developed over a long period of acquaintanceship.

5. Commercial banks operate an informal cartel.

3. Commercial banks as the dominant institutions

This stylized fact both describes the market structure we sometimes assume, and may be explained by our consideration of entry. In the pricing and credit models (Chapters 4 and 7) it limits the players we consider to those who can be described as banks. However, the entry model,

by showing how potential entrants may be induced to leave the market, provides an explanation of the continuing restricted number of competitors on the financial markets. As is typical in LDCs, commercial banks dominate the financial sectors. Central banks appear to be the only institutions with sufficient market (coercive) power to provide effective competitive pressure but, not only are their motives bureaucratic, but their need to take account of the banks' interests limit their freedom of action.

Consider the alternatives for savers. There are stock exchanges in Trinidad and Tobago and Jamaica, but the limited flow of shares and high transactions costs reduce their role as a source of assets for savers. In addition to commercial banks, financial savers have access to life insurance companies, finance companies, trust companies and credit unions. The last of these are significant only in Belize, about half of whose population is rural. Commercial banks held 88% of deposits at Belizean institutions for which data is available (see Luben [1983], Tables 2, 7 and 13). In the Bahamas, commercial banks held 80% of deposits in 1978, while finance companies (who are at the long end of the financial market, like British building societies before the recent financial changes in the U.K.) held 19% [8]. In Barbados, commercial banks had an 86.7% share of total deposit liabilities at deposit-taking institutions[9] while trust companies' share was 12% (trust companies play a role similar to finance companies in the Bahamas, but are largely owned by commercial banks). In Trinidad 42% of the household and unincorporated enterprise (HIE) sector's financial assets were held at commercial banks in 1978 and 13.8% at insurance companies[10].

For borrowers, the picture is rather similar. In the Bahamas commercial banks owned (1978) 66.7% of financial sector assets, while finance companies held 17.9%[11]. For Barbados

---

[8] The deposit-taking institutions included in this comparison are: commercial banks, finance companies, the Post Office Savings Bank , the People's Penny Savings Bank and trust companies. Ramsaran, op.cit., Tables 2.22, 4.3, 4.4, 4.15, 4.13.

[9] The deposit-taking institutions are commercial banks, trust companies and finance companies; inter-institutional deposits have been netted out to avoid double-counting.

[10] These shares are obtained from flow-of-funds data (Bourne [1982]) and while they appear to indicate a more equal distribution of private sector assets, it should be noted that they are more comprehensive than the data given for the other countries and are therefore not directly comparable. Trinidad is the only one of the countries to prepare flow-of-funds accounts.

[11] Ramsaran, op.cit., Table 3.16

commercial bank loans represented 70.6% of credit extended by major financial institutions in 1983. The trust companies (11.4% of total credit), were the major lenders after commercial banks (and are in any case owned by the banks), and insurance companies (7.9%) were an important source of long-term residential mortgages. In Belize commercial banks owned 77.5% of assets of major financial institutions in 1981 (Luben, op.cit., Table 21). Despite their stock market, commercial banks have been the main source of finance for the corporate sector in Trinidad, providing 30.9% of their finance in 1978. Commercial banks held just over 50% of household liabilities.

It seems clear that there are no good substitutes for commercial banks on the side of either the saver or the borrower. Finance and trust companies variously confine their credit to consumer installment or mortgage loans, and for depositors the former can be disproportionately risky. Credit unions require membership. Financial savings with insurance companies are illiquid and their lending (to policy holders) is for mortgages.

4. Bank-customer relations

Caribbean analysts have commented on a link between knowledge, longevity and customer relations (service), especially in the credit context (see also Chapter 7.3). Loan patterns are viewed as having evolved from sugar industry funding and import dependence. Major favoured firms remain concentrated in the retail and wholesale trades, although some of these have diversified into manufacturing. Thus, we have Bourne [1984] commenting:

> "Information costs are the core of loan appraisal costs... (and) will differ sectorally according to the sectoral distribution of prior knowledge and required knowledge."

while Worrell [1985] remarks that:

> "The working capital finance for which firms seek bank loans is provided on the basis of knowledge of the firm and its operations, and local bank managers are the repositories of this information."

Statistical evidence for these associations is not publicly available in the Caribbean but their existence is supported by bank examiners who have detailed confidential knowledge of banks' portfolios. Indeed, four out of seven banks, answering the Central Bank of Barbados' request for

information on their loan and interest rate policy, attached first priority to the characteristics of, or their knowledge and experience of, the individual or company management. This was the only factor common to all the respondent banks.

In addition, the Bank of Nova Scotia's orientation towards consumer credit [12] forced the other banks to compete in this area, a tendency fuelled by rising incomes and close communications with North America. As a result of this, and presumably of its low risk features, banks throughout the region have tended to concentrate a considerable portion of their loan portfolio in personal sector credit.

A relationship resulting from knowledge acquired over a long time period also has effects on banks' deposit policy. Banks in Belize are seen by Barnett [1982] as attaching undue importance to the sectors with which they are well-acquainted, to the point where

"... the banks appear to be competing with each other for existing customers rather than attempting to attract new deposits into the system."

In this context, we may emphasize that interest rates in Belize are not subject to controls. In Barbados banks provide credit facility cover for old customers who take deposits on an informal basis although this directly reduces banks' own revenue since such firms thus avoid bank borrowing. This is a traditional practice arising out of sugar industry trade links, so that the depositors tend to be firms or individuals with commercial or personal links to the companies [13].

5. Collusion

Advocates of liberalization have recognized that banks in LDCs are unlikely to act like perfect competitors. Shaw ([1973], pp.88,123) attributes this to restrictions imposed on entry by the authorities (charter policy), as well as to the effects of administered interest rates. Galbis ([1979], p.41) acknowledges that oligopolistic behaviour will not result in efficiently set rates.

---

[12] "In the Caribbean as in North America, the Bank of Nova Scotia wants to carry the image of the friend to 'the man in the street'. It wants to do this through its highly developed consumer loan program, the Scotia Plan." (Baum [1974], p.77)

[13] Given this specialized market and the risks attached to such deposits, these firms would be ineffective competitors with the banks, though their interest rates tend to be one or two percentage points higher . In 1983 their deposits were 6.7% of total deposits at commercial banks and 50% of trust company deposits. Without detailed information on these firms, it would be double-counting to include such deposits in economy-wide deposits since when not used internally they are held at commercial banks.

With small numbers, oligopolistic behaviour appears indisputable: rational observation on the part of the members of an industry where four firms produce 75% of the output and purchase 75% of the input, [14] would lead them to take account of each other in their actions. Since entry is not exogenously restricted in the Caribbean (see Chapter 6), and we are interested in bank behaviour in the absence of controls, the issue is whether there is some more subtle form of regulation that encourages collusion, or whether tacit collusion[15] is feasible. The well-known difficulties of enforcement make explicit collusion an unsatisfactory explanation. In Chapter 4 we show that, in a search environment with switching costs, *a priori* noncooperative behaviour can give rise to effective collusion. Caribbean economists have generally assumed that their banks do act collusively (see Barnett [1982], p.150; Bourne, [1984], p.33; Worrell [1985], p.62) on the basis of both historical and current observation.

There is historical evidence that banks in the sterling area collaborated in deciding the appropriate response of local rates to movements in the London Bank Rate. Evidence given to the Radcliffe Committee in 1957 by a group of banks describes the process as follows:

> "The normal procedure in the case of currencies firmly linked to sterling is for the local banks to confer among themselves as to the effect the Bank Rate change has on the local structure of money rates."[16]

At this time, the Commonwealth Caribbean was part of the Sterling Area and banks in the region had their head offices, or held their foreign working balances, in London. Collusive behaviour is not surprising in the context of the British banking system's own behaviour: until 1971 British clearing banks operated a cartel for interest rate fixing. It was only the increasing competition from foreign banks and official insistence (Competition and Credit Control) that ended the rate-fixing cartel (Goodhart [1984]). It can be noted that, while the Canadian Act regulating the domestic and foreign activities of banks prohibits any agreement between banks to fix rates of interest, it specifically exempts deposits or loans made or payable outside Canada from the prohi-

---

[14] Four-firm concentration ratio for loans and deposits, Barbados, 1986

[15] By this we mean collusion which is individually rational for the participants.

[16] Radcliffe Committee, *Principal Memoranda of Evidence*, Vol.2, pp.64-65. Quoted in Ramsaran, op.cit., p.125

bition (Baum, op.cit., p.31).

Current observation indicates that collusive behaviour continues in the Caribbean. Banks meet to fix fees, for example; as is evidenced by bank inspectors' observation of the simultaneous imposition of similar fees at all banks, although fees are not immediately observable to banks or customers. In fact, in the Bahamas, the clearing banks' agreement to fix loan and deposit rates appears to be explicit. The prime rate

> "is fixed after discussions among the clearing banks have taken place and the Central
> Bank has given their approval. Other banks then tend to fix their own rates in some
> relationship to the Prime figure operated by the clearing banks." [17]

Since the clearing banks control over 95% of both deposits and loans (in 1978) this agreement effectively determines the market rates. Given this empirical evidence, inference of bank collusion appears well-founded and theoretical support is given by the model of Chapter 4.

3.5: Concluding remarks

In the following chapters we use the stylized facts about institutional structure as the background to models of strategic bank behaviour in stochastic environments and/or those with information differentiated between bank and client. The behavioural 'facts' describe our initial assumptions about bank actions and concerns and motivate our priors about outcomes. We show that, given the assumption that banks have special knowledge of, or acquire information about particular clients, and that their general information is limited, market outcomes are not those financial liberalization theory would lead us to expect.

---

[17] Ramsaran, op.cit., p.141, footnote 81

Table 3.1

Basic Indicators				
	Bahamas	Barbados	Belize	Trinidad
Population 1987 ('000)	240	254	176	1,200
Pop'n Growth Rate	1.55	0.03	2.3	0.7
Life Expectancy (yrs.)	70	75	67	70
Area ('000 sq. km.)	14	.43	23	5
GNP p. cap. 1987 ($US)	10,280	5,350	1,240	4,210
GNP growth 1965-87 (av. % p.a.)	0.9	2.4	1.9	1.3
Average inflation (% p.a.)				
1965-80	6.4	11.2	7.4	14.0
1980-87	6.3	6.1	1.1	6.2
No. of bank offices	52	49	16	112
Pop'n per bank office	4,500	5,102	9,085	10,077
Sectoral contribs. to GDP (%) Aver.1980–84*				
Agriculture	4	6.8	21.8	3.0
Industrial Activity				
Total	11	14.0	11.8	37.4
Manufacturing	n.v	10.6	10.2	8.8
Construction	2	6.4	5.4	14.6
Wholesale & Retail Trade	23	28.4	15.2	12.2
Transp. & Communic.	10	6.4	7.0	9.6
Gov't & Services	n.v	12.4	11.4	12.2

NOTES:

Sources: World Bank [1989] *World Development Report*, Oxford University Press; United Nations [1988] *National Accounts Statistics: Analysis of Main Aggregates, 1985*, United Nations Publications

* Except for The Bahamas for which only 1978 data was available.

n.v denotes not available

In the *Sectoral Contributions to GDP* Agriculture includes agriculture, hunting, forestry and fishing; Industrial Activity includes mining and quarrying, electricity, gas and water, in addition to manufacturing; Wholesale and Retail Trade includes distributive trades, restaurants and hotels; Government and Services includes finance, insurance, real estate and business services, community, social and personal services.

BAHAMAS: Bank office data from 1979; Source: Ramsaran [1984].

BARBADOS: Bank office data from 1984; Source: Central Bank of Barbados

BELIZE: Bank office data from 1983; Source: Luben [1983]

TRINIDAD & TOBAGO: Bank office data from 1982; Source: Bourne [1984]

Table 3.2

Monetary & International Transactions: A Snapshot Comparison, 1982* Ratios to GDP at Current Prices (percentage)				
	Bahamas	Barbados	Belize	Trinidad
Monetary Survey				
Net Foreign Assets	1.3	0.2	-9.1	41.7
Dom. Credit to Gov't (net)	7.0	8.4	20.6	-19.8
Dom. Credit to Priv. Sector	33.3	38.1	31.6	33.8
Quasi-Money	27.2	29.6	28.5	27.6
Tot. Monetary Liabilities	38.2	42.2	42.1	42.1
Other Financial Instits.				
Net Foreign Assets	6.4	...	...	0.3
Claims on Gov't	0.4	-	...	0.8
Claims on Priv. Sector	9.8	6.4	...	14.5
Deposits	9.6	4.4	...	9.7
Liquid Liabilities	47.3	46.6	...	51.4
International				
Exports, fob	169.2	26.0	50.7	42.0
Sugar exports	-	3.1	18.0	0.3
Citrus exports	-	-	5.3	-
Petroleum exports	-	-	-	36.6
Imports, cif	211.7	55.7	...	50.5
Current Acct. Balance	-4.4	-1.8	-4.1	-4.7
Current GDP, 1982* ($US mn.)	1449	990	193	7907

Sources: IMF, *International Financial Statistics, 1987 Yearbook*; United Nations [1988] *National Accounts Statistics: Analysis of Main Aggregates, 1985*, United Nations Publications

* Data for Belize is for 1984; ... denotes not available; - not applicable or negligible

The *Monetary Survey* consolidates the accounts of the monetary authorities and deposit money banks, measuring money creation.
- Net Foreign Assets are the sum of foreign assets less the sum of foreign liabilities of the title institutions.
- Domestic Credit to Government is the sum of claims on government less the sum of government deposits of the title institutions.
- Quasi-Money are the time, savings and foreign currency deposits of residents with the monetary authorities and deposit money banks. Total Monetary Liabilities are the sum of quasi-money, currency outside banks and private sector demand deposits, i.e. M2

Other Financial Institutions include institutions such as savings and mortgage loan institutions, post office savings, development banks, building and loan associations and life insurance companies. Their deposits are time and savings deposits unless otherwise noted in country notes.

Liquid Liabilities consolidate total monetary liabilities and deposits at other financial institutions, eliminating intra-sectional items. It is a broader measure of M3.

BAHAMAS: Commercial bank data (included in the monetary survey) relate to accounts of the authorized dealers which are permitted to deal in gold and all foreign currencies. Other financial institutions consolidates the accounts of licensed banks and trust companies permitted to undertake domestic business, other than commercial banks. Their deposits include foreign currency deposits.
BARBADOS: Other financial institutions are the trust companies.
BELIZE: Belize's data is for 1984. No data on other financial institutions is available.
TRINIDAD & TOBAGO: Monetary Survey domestic credit to the private sector is the total for local government, business and individuals. Other financial institutions consolidates the accounts of banking institutions and life insurance companies.

Chapter 4: Deposit Rate Determination

4.1: Introduction

Since the 1973 publication of Shaw's and McKinnon's books, analyses of finance in LDCs have stressed the role commercial banks' deposits play in mobilising savings. These analyses argue that, unless constrained by regulation, competitive commercial banks are led by available loan opportunities to set deposit rates at the high levels which reflect the opportunity cost of capital in poor economies. This claim has been extensively modelled and tested at a macroeconomic level, the emphasis being on the interest elasticity of aggregate savings. However, commercial bank deposit rate determination in the face of financial savings behaviour has not been analyzed. This chapter aims to do so.

The chapter makes four points. First, we show that when depositors value the intermediary services provided by banks, and both bank and depositor have private information, there are utility gains to the depositor of remaining with a bank. That is, the depositor has switching costs. This implies a reduction in deposit supply elasticity, which encourages tacit price collusion by banks. The resulting monopsonistic deposit rate can frustrate savings mobilisation following liberalization and makes deposit rate regulation an appropriate policy.

Second, the modelling device of switching costs has facilitated a growing number of analyses which examine repeated, stable relationships between buyer and seller. All of these models impose the switching cost exogenously. We derive switching costs endogenously as a tradeoff between service quality and the interest rate resulting from relationship-specific quality enhancement in, the long-run of a market where customers must search for quality. On this basis, we draw a distinction between switching costs and transactions costs, arguing that it is because the former entail *ex post* matching that they can be an important source of monopoly power. The distinction is of general interest because it can indicate markets in which price competition is less likely.

Third, we believe that switching costs can provide an explanation of the durable bank-customer relationship much discussed in the banking firm literature.

Fourth, the model provides a rationale for the collusive behaviour Caribbean economists have argued characterizes their banking sectors. Such behaviour seems at odds with the instability of cartel arrangements, but we show that there are conditions under which collusion is non-cooperatively sustainable.

The chapter is structured as follows. Pricing games are usually able to consider markets for distinct goods. The substitutability between domestic bank deposits and other financial assets could raise doubts about the validity of a similar approach in the financial context. Section 4.2 describes the institutional features which justify this treatment. Similarly, the usual analysis of financial deepening treats deposit flow as money demand and regards bank deposits as homogeneous. Section 4.3 explains our view of quality differentiated deposit accounts and discusses the similarities between our switching cost hypothesis and the bank-customer relationship analyzed in the banking literature. Section 4.4 describes and solves the model of depositor behaviour, demonstrating the development of switching costs. These are used in our deposit market application of Klemperer's [1987a] model to show, in Section 4.5, that tacit collusion can be a noncooperative equilibrium.

There is a point worth stressing here. Although our analysis concludes, in contrast to the liberalization hypothesis, that post-regulation banks will collude on the deposit rate, holding it beneath the social opportunity costs of funds, the priors of our analysis do not diverge in substance from those of, for example, Shaw. We have however explicitly considered consumer choice of bank services in a situation of imperfect information, and bank strategic response to such behaviour. This is significant in two respects. From the descriptive point of view, it suggests that liberalization theory did not go sufficiently far in analyzing the choice-theoretic implications of their own postulates, and from the policy viewpoint it suggests that the financial deepening thesis may have been misleading in its inference that no regulation is the optimal policy stance.

Thus, following discussion of our results in Section 4.6, we go on in Section 4.7 to show that, in markets of the type described here, regulation is the appropriate policy response.

4.2: Institutional background and stylized facts

Two stylized facts from the Caribbean area allow us to isolate a domestic commercial banking sector with a limited number of firms as the only financial savings repository available to depositors (savers):

Insulated domestic financial markets

In a truly open financial market, international arbitrage implies that financial prices and services would be determined at world levels. Exchange controls and nonconvertible currencies insulate the domestic financial markets from this possibility. The empirical justification of this 'fact' has been described at some length in Chapter 3.4. Here we summarize the key points.

As a result of exchange control regulations, domestic financial institutions do not in general compete (legally) with foreign intermediaries for domestic deposits. Residents can only borrow or invest abroad with exchange control permission which is given only after consideration of each specific case. The non-convertibility of regional currencies reinforces the insulating effect for the small depositor although, in at least the Bahamas and Barbados, the relatively widespread availability of US dollars through the tourist industry, does mean that residents have access to foreign currency outside legal channels. However, as we concluded in Chapter 3, the balance of the evidence suggests that legal risks and transactions costs effectively prohibit the illegal transfer of funds abroad for those residents who do not have access to foreign currency deposits.

A small number of banks with no effective competition from alternative institutions

In the absence of this 'fact', the consumer deposit decision would have to reflect the alternative services and (adjusted) rates of return offered by different types of institutions, and the analysis of bank interest-rate setting would have to take account of strategic action by issuers of alternative liabilities. A model attempting to capture this degree of reality would rapidly become intractable. The acceptability of the simplification basically depends on whether or not it is reasonable to take the view that liabilities of other financial institutions are not seen as substitutes by depositors, and that banks are aware of this and act as market leaders in deposit-rate setting. This view is based on the following observations.

While there are stock exchanges in Trinidad (and Jamaica), the limited flow of shares and high transactions costs would reduce their role as a source of assets for savers. Therefore, on the savings side of the market, potential competitors to commercial banks consist chiefly of insurance companies, finance companies and credit unions, with their relative importance varying according to the economic and geographic features of the different countries (see Table 4.1). In Belize, for example, where there is a scattered rural population, credit unions are relatively important with members' shares at the major credit union representing 8.4% (in 1980) of aggregated deposits at commercial banks, the Government Savings Bank and the credit union itself. Commercial banks held 88% of these deposits [1]. In Barbados, life insurance companies probably represent the main alternative repository for financial savings, if trust companies are excluded - in the absence of flow-of-funds data a direct comparison is not possible. At the end of July 1986, commercial banks had an 85.9% share of total deposit liabilities at deposit-taking institutions[2], while trust companies' share was 12%. Since trust companies were established by commercial banks to handle the longer-term components of their portfolios, they are unlikely to be price competitors, and for our purposes here could legitimately be viewed as aggregated with the banks. The significance of life insurance companies is indicated by premiums of BDS$34.8 million in 1984, compared to an increase of BDS$104.6 million in commercial bank deposits and BDS$5.0 million in trust company deposits[3]. However, insurance company liabilities are a contractual obligation, with no option (if we exclude the costly one of policy cancellation) to withdraw depending on the interest rate or service received.

Insurance companies are also an important feature of the savings market in the Bahamas where 20 companies received B$17.6 million in life premiums in 1977 while total deposits at commercial banks rose by B$22.7 million[4]. Commercial banks held 80% of deposits at deposit-taking institutions in 1978 with the remainder virtually all held by finance companies since other

---

[1] Luben [1983], Tables 2, 7, 13

[2] The deposit-taking institutions are commercial banks, trust companies and finance companies; inter-institutional deposits have been netted out to avoid double-counting.

[3] See Central Bank of Barbados [1988]

[4] Ramsaran [1984],Tables 3.7 and 2.2; the figures refer only to Bahamian dollar business.

institutions' share was negligible[5]. Trust companies in the Bahamas, even the resident [6] ones, exist primarily for the external business generated by the Bahamas' tax haven status. The finance companies play a role similar to trust companies in Barbados, mainly funding mortgages through deposits. They are, with perhaps one exception, independent of the commercial banks.

Flow-of-funds data for Trinidad and Tobago [7] (reproduced in Bourne [1982]) indicate that commercial banks' deposits remain dominant in private financial assets. Forty-two percent of the household and unincorporated enterprise (HIE) sector's financial assets was held at commercial banks in 1978 and 13.8% at insurance companies[8]. Finance houses represent a potential near-substitute for banks in Trinidad (and in Barbados) where they offer, for example, chequing facilities. However, the risk attached to their deposits is likely to deter most depositors, and is evidenced by several recent bankruptcies (see Parris [1984]). Trinidad and Tobago is the only one of the countries with a long-established stock exchange but shares were only 9.5% (Bourne [op. cit.], Table 7) of instruments in HIE financial assets in 1978.

To summarize: the limited availability of shares and the high transactions costs of trading in them preclude their consideration. None of the other institutions described would appear to present a challenge to commercial banks' price setting. Financial savings with insurance companies are relatively illiquid. Trust companies and finance houses may be either closely connected by ownership with the banks, considered highly risky by the depositing public, or offer services sufficiently different from banks as to neutralize their rivalry or place them in a different segment of the market. Credit union-type arrangements usually require specific geographic or professional affiliation.

---

[5] The deposit-taking institutions included in this comparison are: commercial banks, finance companies, the Post Office Savings Bank, the People's Penny Savings Bank and trust companies. Ramsaran [op.cit.], Tables 2.22, 4.3, 4.4, 4.15, 4.13.

[6] Resident financial institutions are those with a licence to carry on Bahamian dollar business, that is, they are resident for balance of payments and exchange control purposes.

[7] Trinidad and Tobago is the only country considered for which flow of funds date is available.

[8] While these shares appear to indicate a more equal distribution of private sector assets, it should be noted that they are more comprehensive than the data given for the other countries and are therefore not directly comparable.

Table 4.1

Institution	Bahamas (1978)	Barbados (1983)	Belize (1981)	Trinidad (1978)
Institutional Distribution of Deposits (percentage)				
Commercial Banks	80.3	86.7	88.6	42.0
Finance Cos.	118.6	1.8	-	sc
Insurance Cos.	...	...	...	13.8
Trust Cos.	~	10.6	-	sc
Savings Banks	1.1	-	3.0	sc
Public Fin. Insts.	...	...	...	9.2
Credit Unions	-	0.9	8.4	sc
Pension Funds	...	...	...	9.2
Other	-	-	-	26.0

NOTES:

- denotes not applicable; ... denotes not available; ¯ denotes negligible sc = see country note

As far as possible where relevant, inter-institutional deposits have been netted out to avoid double-counting

BAHAMAS: Bahamian dollar assets. Source: Tables 2.22, 4.2-4.4, 4.13 in Ramsaran [1984]

BARBADOS: Source: Central Bank of Barbados, *Annual Statistical Digest, 1986*

Data is end-year, except for credit unions which are at 31.03.83

BELIZE: Proportional distribution of the deposits at the listed institutions.

Source: Tables 2, 7, 13 in Luben [1983]

TRINIDAD & TOBAGO: Financial assets of households and unincorporated enterprises (HIE). As this is flow-of-funds data, the coverage is more comprehensive than that available for the other countries and is not directly comparable. Furthermore, by definition it excludes government and corporate assets.

sc indicates that this category is included in 'other'.

Included in 'Other' are the central bank, building societies, credit unions, the local corporate sector, trust companies, finance houses, central and local government, trust companies and the foreign corporate sector.

Source: Table 6 in Bourne [1982]

4.3: Switching Costs in the Banking Market

Switching costs are usually said to occur when a consumer finds it costly to start consuming, or to switch suppliers of, a good. The endogenous development of switching costs obtained in Section 4.4 is based on potential switching between deposit accounts which provide vertically differentiated services; such services are imperfectly observable prior to account opening and customer-specific once an account has been opened. Our notion of switching costs has precedents in the literature on the banking firm which describes the basic switching cost result, that clients remain loyal to past suppliers, as the bank-customer relationship.

Interest in the effects of switching costs in product markets has recently been reawakened by Klemperer ([1986a], etc.)[9]. Since it is costly for a consumer to switch suppliers, demand becomes less price elastic: instead of a smooth flow of consumers between suppliers in response to marginal price differentials, consumers switch only in response to prices low enough to compensate for switching costs. Klemperer identifies four sources of switching costs: transactions costs, learning costs, artificial costs such as repeat purchase coupons, and psychological costs such as brand loyalty and/or habit. In the models he develops to demonstrate how switching costs affect firms' strategic behaviour, Klemperer models switching (or start-up) costs as an exogenous cost arising when a consumer buys from a firm from which it did not purchase in the previous period. In our Section 4.4 switching costs arise endogenously in the sense that they are given by the interest rate premium required to induce a customer anticipating improved service to switch banks, that is, as a utility tradeoff between the rate of interest and service quality in long-run equilibrium.

We begin in this section by describing our concept of vertically differentiated deposit accounts, preview our hypothesis of switching costs development and note the parallels and differences between the loyalty we describe and that discussed in banking models.

---

[9] See Klemperer, op. cit., and Section 4.6, for references to several other models incorporating the effects of switching costs. The switching costs description closest to our own is that of Schmalensee [1982], where consumers' knowledge of a 'pioneering' brand, and uncertainty about the quality of the entrants' products, gives them a high opportunity cost of experimenting with another brand.

Vertically Differentiated Deposit Accounts

Monetary and financial analyses treat deposits as homogeneous flows in response to the deposit rate. We view deposits at different banks as heterogeneous goods. Consumers save via an intermediary both for the interest income they earn and for the intertemporal and spatial transactions facilities offered by the intermediary. This is recognized by the financial liberalization literature when it discusses, for example, the payments mechanism provided by banks. In this sense, the counterpart in the bank market to the differentiated product in the goods market is not the deposit flow *per se* but the deposit account. In opening an account, the consumer is not indifferent either as to the quality of service he receives, the variety of services available to him or the institution he banks with. This is clear in the last case, where some institutions may be considered more risky than others. With regard to the former two, different depositors will have different service requirements but there are also services which all consumers would value equally. Thus, deposit accounts combine elements of both horizontal and vertical differentiation.

In the product market, heterogeneity is usually classified as horizontal or vertical differentiation, corresponding to the common distinction between goods that differ because they meet different needs/tastes (horizontal) and goods that differ in quality (vertical). Products are said to be horizontally differentiated when, if sold at the same price, different consumers would choose to purchase different goods because they prefer one specification rather than another, i.e. their tastes differ. With vertically differentiated products, on the other hand, there is a unanimous ranking of the goods by consumers, so that, if sold at the same price, all consumers would choose to purchase the good they judge to be "better" - that of higher quality, with a higher quantity of all the common characteristics (Geroski, Phlips and Ulph [1985]). We denote the intertemporal and spatial transactions facilities offered on each deposit account as service quality. It is differences in the degree to which facilities are provided that link banks' roles as asset transformer, monitor and assessor, inducing loyalty. Hence our focus on vertical differentiation.

Observed features of the banking markets in Caribbean economies indicate that different deposit account holders at the same bank do receive differentiated service (even when the

accounts have similar interest rates and maturity), and ownership of an account at one bank rather than another does imply that customers have access to different services. Some banks pay interest on chequing accounts, interest penalties on early withdrawal of time deposit accounts vary between banks, and are negotiable for certain customers. Credit card and cheque guarantee cards are offered to some customers by a subset of the banks. The availability of financial advice services differs between banks and among customers. Different banks are more or less willing to offer loan facilities to depositors, depending on their criteria and customer evaluation. The number of branches and queue length varies between banks.

There is a similarity between this treatment and Salop's [1977] description of the bank's product as a service package made up of a number of attributes and convenience facilities, with consumers differing in their needs and valuations of services (a horizontally differentiated package)[10]. His view of the bank product as a service package has the advantage of encompassing both the asset and liability sides of the bank portfolio. However, our emphasis on the deposit account is more appropriate to our hypothesis that the repeated and ongoing nature of the banking transaction is crucial in determining the nature of price competition. The deposit account is the usual means by which buyer and seller come into contact on the banking market: a bank customer normally has to open an account on the liability side of the bank's portfolio, whether current or interest-bearing, before loans and other bank services become available to him. This is fundamental to the intermediary role - not only does the bank intermediate between separate borrowers and lenders but, partly in order to do so, it also intermediates between the intertemporal requirements of the same economic unit, as McKinnon's [1973] stylized model of the complementarity of financial saving and capital for the saver-investor stressed.

Like Salop, most theoretical models of differentiated products consider only one of the two descriptions. In reality, it is obvious that products and services combine both kinds of differentiation and it is not always clear in which category a particular attribute should be placed.[11] We

---

[10] The attributes Salop considers are minimum balances required for free chequing, activity charges related to minimum or average account balances, financial counselling, safe deposit boxes, direct deposits, third party payments, overdraft and line-of-credit facilities, while convenience facilities include bank hours, the number of branches, drive-in windows and average queue lengths at teller windows.

[11] Ireland (Chapter 7, [1987]) considers a model where a product may be both horizontally and vertically

argue that, while the deposit account is both vertically and horizontally differentiated, it is the time varying quality dimension that provides depositors with an incentive to stay and therefore gives rise to switching costs. The key assumption is that all depositors view the ability to trade consumption intertemporally as an important quality dimension of bank service. Thus the implicit commitment (subject to certain conditions) to future overdraft, line-of-credit and other loan facilities is a service offered by banks, and the more likely does an account appear to provide future credit, the better it is judged by *all* consumers. Similarly, the courtesy and consideration with which a customer is treated can evidently be considered as a quality attribute.[12]

Even when we consider those services which appear best classified as horizontally differentiated, they often include a quality dimension, similar (i.e. time-varying, customer-specific as a result of imperfect information) to that of characteristics which, at the same price, would be equally ranked by all depositors. For example, while interest payments on chequing accounts and cheque guarantee cards are of value only to that subset of depositors who wish to hold a chequing account, guarantee cards (and credit card facilities) are only issued to customers after a bank has acquired some knowledge of them[13].

Our hypothesis is summarized as follows. Individuals and firms save in the form of financial deposits with banks both because they earn a rate of return to compensate them for foregone present consumption, and because banks provide them with certain services which include the payments mechanism, financial intermediation, security etc. Among the most important of these services is that deriving from a bank's role as financial intermediary between borrowers and lenders. As such, the bank also acts as an intertemporal intermediary for the individual economic unit, allowing it to smooth consumption over time and in different states of nature (and even as a

---

differentiated, with either one of the differentiates having a dominant role.

[12] It can also be argued that the branch network and average queue length are quality differentiates, since all customers would prefer banks with shorter queue length and more branches, if paying the same rate of interest. This is evident for the personal sector, and to the extent that businesses have several outlets and have to visit banks, we would also expect it to be true for the business sector. However, neither feature is important in the present context since they are ascertained at low (negligible) cost, and are not changing in a customer-specific fashion over time.

[13] Other horizontally differentiated services are financial advice services addressed to particular economic sectors (such as the agricultural), safe deposit boxes and third party payments.

spatial intermediary via the payments mechanism, with the issue of cheque guarantee cards, third party payments etc.). But both banks and customers are heterogeneous and operate in a situation of imperfect information which affects their profit and utility maximising behaviour. Furthermore, they mutually recognize some of the results of this imperfect information. Customers cannot perfectly observe the quality of the intermediary service offered by banks: different banks may have different criteria in offering intermediary service and may be more or less risk averse. The courtesy with which a customer is treated cannot be observed prior to depositing. In Nelson's widely-referred-to terminology, service quality is an experience good. Thus customers must search and wait in order to find satisfactory service. In addition, customers recognise that the quality of the intertemporal intermediary service - the degree to which banks are willing to lend to depositors and to facilitate their payments - depends on the bank's acquiring information about the customer. Because customers differ in their risk characteristics [14], a profit-maximizing bank must try to differentiate among them and one method of doing so is inference from the deposit behaviour of the customer. The more stable the deposit behaviour of the customer, the more able the bank is to provide this service. Thus the customer expects the service he receives to improve over time and has an incentive to build a reputation with the bank. As a result the customer develops switching costs: once he has started to deposit with a bank the expectation that the reliability of his behaviour will improve the service received next period locks him into that bank.

The strength of these factors would be expected to be much greater in underdeveloped financial markets where economic units have poor access to other sources of credit and information systems such as credit ratings have not been established. This may account for the fact that the deposit account has not been considered as a means of sorting customers in information-oriented analyses of credit markets. In addition, in small markets a bank seems far more likely to gain valuable information through a deposit account which will tend to reflect the transactions of the depositor more completely than in the case where the depositor may be using a variety of

---

[14] The possible adverse selection effects on credit markets have been analysed by (principally) Stiglitz and Weiss in several well-known papers, [1981] for example. This issue will be considered in Chapter 7.

financial intermediaries. This does not preclude the depositor having access to other sources of credit or the bank having other means of sorting customers. What is claimed is that (especially) in LDCs, this factor is important enough to give rise to substantial switching costs.

The Bank-Customer Relation

The notion of switching costs recalls the bank-customer relationship discussed in the banking literature, usually with reference to credit rationing. Hodgeman argued that customers' joint demand for loans and revenue earning bank services meant that prime customers (those with the largest deposits) are more valuable to the bank than the current loan would indicate (see Blackwell and Santomero [1982] and Santomero [1984]). They are therefore less likely to be rationed. Wood [1974] stressed the time-dependent nature of the relationship: *ceteris paribus* , the bank would prefer a loan applicant more likely to maintain a long-term relationship, and the depositor would expect more favourable loan terms as a result of maintaining a sizeable account balance.

These descriptions do not explain how a durable relationship would be developed and maintained and possible effects on the deposit supply itself do not appear to have been analysed. Wood stresses deposit *size* rather than information gained over time as the crucial factor. This is not compelling for two reasons: if size of deposit is the bank's main interest, the depositor could obtain similar benefits by transferring large deposits to another bank rather than maintaining a relationship with one bank, i.e. size alone is insufficient to account for durability. Secondly, both descriptions emphasize (large) business customers, since these are the customers most likely to have substantial deposits. However, commercial banks are evidently also interested in personal customers and in small, profitable business customers. It is the aggregate deposits of the former group that provide banks with funding for their revenue-making loans. The size alone of such deposits would not be sufficient to account for durability. Nor does the phenomenon appear to have been generally linked with informational considerations. In Santomero's [1984] survey of the banking firm literature, the discussion of credit rationing treats the customer relation and information problems in separate subsections.

Flannery [1982] introduces informational and learning considerations into banking relation-ships in the retail deposit market. Retail deposits are viewed as a quasi-fixed factor of production for banks, in that the bank has sunk set-up costs in establishing the deposit account. On the con-sumer side the costs arise from learning alternative investments, new procedures etc. The value of this information endures beyond the initial transaction and cannot be transferred to another bank. Thus, both bank and customer have transaction-specific investments, giving them an incen-tive to continue the relationship in order to amortize setup cost. Banks therefore price retail deposits above the marginal cost of funds in the short run to keep depositors. Neither setup costs nor higher short-term rates seem likely to create durability since a future lucrative offer could tempt the depositor away. Devinney [1986] also recognises that mutual imperfect information is at the root of the bank-customer relationship but analyses it only in the credit rationing context. The bank updates information about the customer through granting a loan and observing repay-ment ability.

The banking literature recognizes the durability of the bank-customer trading relationship but does not explain it. Explanations on the deposit side tend to focus on wealth which is transferable and those based on credit ignore the deposit account as a logically prior (and less costly) source of information. Recognised as such by the customer, who must trade the expected loss in service against the gain in interest income if he leaves in response to an interest differen-tial, it gives rise to the switching costs that are the source of durability. This result is formally derived in a model of optimal dynamic savings behaviour in the next section.

4.4: The Depositor's Problem

Introduction

In the following model we combine two applications of dynamic programming in problems of optimal sequential choice under uncertainty by adding uncertain service quality to a utility function defined on deterministic consumption. The consumer simultaneously searches for satis-factory service at banks and makes optimal intertemporal consumption decisions which depend

on the rate of interest. The search for service is analogous to the sequential information gathering about price or wage usually modelled in search theory - see, for example, Kohn and Shavell [1974] or Lippman and McCall [1976]. A series of papers - see Phelps [1962], Levhari and Srinivasan [1969], Samuelson [1969] and Mirrlees [1974] - applied dynamic programming to lifetime saving and portfolio selection when the value of wealth, income or the rate of return is stochastic. Consumption in these models is therefore uncertain while in our model only service is uncertain, it being assumed that consumers learn the deposit rate costlessly.

In analyses of a competitive or monopolistic product market, when the quality of a good is imperfectly unobservable prior to purchase, and consumers differ in their willingness to pay for quality, it is often assumed that the consumer may use product price as an indicator of quality prior to purchase (see, for example, von Weizsacker [1980], Shapiro [1982], Wolinsky [1983]), equilibrium price incorporating a quality premium sufficient to induce the seller to produce the quality expected at that price. These models require that the consumer have some firm-specific information about quality prior to purchase. A similar inference procedure is not attractive in an oligopolistic deposit market where quality is customer/bank-specific rather than firm-specific; each bank offers a range of qualities (here) encountered randomly by the customer. Hence, in this analysis, the depositor is assumed to judge quality only on the basis of his assessment after a bank visit.

We proceed by deriving the optimum consumption and service choices of the depositor and then consider the implications of these choices for depositor behaviour by the use of comparative statics, by consideration of reservation service quality and by examining the utility interaction between service quality and the rate of interest.

The model

We consider an infinitely lived[15] consumer whose objectives are to maximize the sum of discounted utilities of consumption, q_t, subject to an intertemporal budget constraint, and to find

---

[15] While the idea of an infinitely lived consumer is unrealistic, it simplifies the problem, and does not appear too great a travesty when we consider that individuals or firms seldom behave as if they have a terminal date in mind.

a satisfactory level of bank service quality.[16] All savings are held in the form of bank deposits. In any single period the depositor has the utility function:

$$U(q,s) = \frac{q_t^{1+\alpha}}{1+\alpha} + s, \qquad \alpha < 0 \qquad (4.1)$$

which he maximizes in all periods subject to the intertemporal budget constraint:

$$W_t = (W_{t-1} + y - q_{t-1})(1 + r), \quad all\ t \qquad (4.2)$$

and the following boundary conditions: initial deposits (wealth) are W_0 , and $\lim_{t \to \infty} W_t = 0$. q_t is consumption, $-\alpha$ is the constant intertemporal elasticity of the marginal utility of consumption, s is the level of service quality received, W_t are deposits (wealth) in period t, y is constant income received each period, and r is the exogenous deposit rate of interest paid each period, assumed known to the depositor at zero cost prior to depositing: the deposit rate is posted in the bank, advertised or ascertainable by a phone call. Service quality at any bank is not public information and a depositor can only evaluate service at the bank he has visited after a period as customer; the depositor must therefore search among banks for satisfactory service. His reservation service level is denoted s^*.

Over his 'infinite' lifetime the depositor maximizes the time separable utility function:

$$V = \sum_{t=0}^{\infty} \frac{1}{(1+\rho)^t} \frac{q_t^{1+\alpha}}{1+\alpha} + \sum_{t=0}^{\infty} \frac{1}{(1+\rho)^t} E s_t \qquad (4.3)$$

where ρ is his rate of time preference. E is the expectations operator. For the depositor searching for satisfactory service, service quality is a random variable with cumulative distribution function, F and probability density function $f(s)$. The individual's consumption occurs at discrete equal periods of time (each period of time being considered as unity). The state of the system is described by W_t. At the beginning of t the consumer decides to consume q_t of the sum of these deposits and his constant income, y . The amount of deposits available for consumption in the

---

[16] The maximization problem can also be taken as representative of a firm, if we take q as retained earnings, and abstract from the firm's profit maximization problem, redefining the wealth and income variables in the budget constraint as initial equity and profit, respectively. Since the results of the model follow from the treatment of the service variable, explicit consideration of the firm does not appear to add much.

next period is therefore $W_{t+1} = (W_t + y - q_t)(1 + r)$.

This is a multistage problem, where in each period the individual has another multiperiod problem with one less period and different initial deposits. His decision sequence being the choice of q_t which determines the state W_t in the next period. The time separability of the lifetime utility function means that the current decision depends only on the current state W_t and income. Dynamic programming principles are commonly used to solve this class of problems.

The dynamic programming recurrence relation (see (4.4) and (4.5)) allows us to treat the multistage problem as a single maximization problem. It reflects Bellman's Principle of Optimality:

> "An optimal policy has the property that, whatever the initial state and initial decision [W_0 and q_0 in our case] are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."[17]

Thus, the maximum utility from state W is found by maximizing the sum of the current single period utility, $\frac{q_t^{1+\alpha}}{1+\alpha}$, plus the maximum utility of consuming from the resulting next state (W_{t+1}) onwards. This latter quantity is V_{t+1}. Since we are considering an infinite time horizon, the additive utility function and constant ρ implies that preferences for consumption in any period are independent of the period (the age of the consumer), and the budget constraint is independent of period: the environment is stationary so that $V_{t+1} = V_t$ all t.

Consider a depositor who visits banks and opens a deposit account. Let $V(W,s)$ be the depositor's discounted lifetime expected utility from following his optimal strategy. $V(W,s)$ depends on whether service has been found at or above the cut-off level, s^*, or whether the depositor is still searching for such service. s^* is the same in every period, provided there are no changes in the options available to the depositor.

For the depositor who has so far encountered unsatisfactory service, $s < s^*$, the search continues and $V(W,s) \equiv V^B(W)$, the optimized expected value of utility when a deposit has been made but before quality is observed. By the recurrence relation of dynamic programming $V^B(W)$

---

[17] Quoted in Larson and Casti [1978]

is defined by:

$$V^B(W) = \max_{q,s} \cdot \left\{ \frac{q_t^{1+\alpha}}{1+\alpha} + Es + \frac{1}{1+\rho} E\left[V(W,s)\right] \right\}$$ (4.4)

where W is given by (4.2). That is, the expected maximal utility for the searching depositor is maximum consumption utility and expected service this period, plus the discounted future utility of pursuing an optimal policy, given the wealth determined by his optimal choice this period and his knowledge of the market distribution of service.

Having observed service above the reservation level, $s \geq s^*$, $V(W,s) \equiv V^G(W,s)$, where the recurrence relation implies:

$$V^G(W,s) = \max_{q,s} \cdot \left\{ \frac{q_t^{1+\alpha}}{1+\alpha} + s + \frac{1}{1+\rho}\left[V^G(W,s)\right] \right\}$$ (4.5)

The optimized value of utility for the satisfied depositor is his optimal consumption and the identified service in the first period, plus the discounted future utility from pursuing an optimal policy, given the wealth determined by his optimal choice this period and service quality expected, where expected service quality reflects his discovery of at least his reservation service level at his current bank.

Finding $V^G(W,s)$

$V^G(W,s)$ is the expected value of the discounted stream of utility of consumption and service generated by pursuing an optimal policy from the current period onwards, given that the consumer has found at least satisfactory service quality. From its (4.5) definition, $V^G(W,s)$ is additively separable in q and s, linear in s, and defined for each service quality, $s \geq s^*$. We may therefore analyse its components separately, finding optimal q for fixed s received. We write $V^G(W,s) = v(W) + v^G(s)$, where $v(W)$ is the discounted utility value of consumption for fixed s and $v^G(s)$ is the expected value of the discounted stream of utility from service of at least reservation level. $v^G(s)$ varies according to whether the depositor expects service quality to be stationary or increasing. Both cases are considered below. We begin by considering the optimal choice of q.

Optimal consumption, q

To determine optimal q, the consumer's problem is treated as an optimal control problem in discrete time where q_t is the control and W_t is the state variable determined by the choice of q_{t-1}. The problem is to maximize the time separable lifetime utility function:

$$\sum_{t=0}^{t=\infty} \frac{1}{1+\rho^t} \frac{q_t^{1+\alpha}}{1+\alpha} \tag{4.6}$$

subject to (4.2), and it is assumed that consumption takes place at the beginning of the period. Define $v(W)$ as (4.6) evaluated at the optimal value of q.

Using the Lagrange multiplier approach, let $\frac{\lambda_t}{(1+\rho)^t}$ be the multiplier/costate variable. Then, from (4.2) and (4.6):

$$L = \sum_{t=0}^{t=\infty} \frac{1}{(1+\rho)^t} \left[\frac{q_t^{1+\alpha}}{(1+\alpha)} + \lambda_t \left\{ (W_{t-1} + y - q_{t-1})(1+r) - W_t \right\} \right] \tag{4.7}$$

The first order conditions are:

$$\frac{\partial L}{\partial q_t} = q_t^{\alpha} - \lambda_{t+1} \left[\frac{1+r}{1+\rho} \right] = 0, \quad all \ t \tag{4.8}$$

$$\frac{\partial L}{\partial W_t} = \lambda_{t+1} \left[\frac{1+r}{1+\rho} \right] - \lambda_t = 0, \quad all \ t \tag{4.9}$$

$$\frac{\partial L}{\partial \lambda_t} = (W_{t-1} + y - q_{t-1})(1+r) - W_t = 0, \quad all \ t \tag{4.10}$$

$v(W)$ is found by simultaneous solution of these three FOCs. The first order homogeneous difference equation in (4.9) gives:

$$\lambda_t = \left[\frac{1+\rho}{1+r} \right]^t \lambda_0 \tag{4.11}$$

where λ_0 is the initial value of the costate variable. Substituting (4.11) in (4.8):

$$q_t = \left[\frac{1+\rho}{1+r} \right]^{\frac{t}{\alpha}} \Lambda^{\frac{1}{\alpha}} \tag{4.12}$$

where Λ is an arbitrary constant including λ_0. Raising q_t from (4.12) to the power $(1+\alpha)$, and dividing by $(1+\alpha)(1+\rho)^t$, gives

$$\frac{q_t^{1+\alpha}}{(1+\alpha)(1+\rho)^t} = \Lambda^{\frac{1+\alpha}{\alpha}} \frac{\left[\frac{1+\rho}{1+r}\right]^{\frac{t(1+\alpha)}{\alpha}}}{(1+\alpha)(1+\rho)^t}$$

$$= \frac{\Lambda^{\frac{1+\alpha}{\alpha}}}{1+\alpha} \left[\frac{1+\rho}{(1+r)^{1+\alpha}}\right]^{\frac{t}{\alpha}}$$

for each period t. Therefore, summing over t

$$\sum_{t=0}^{\infty} \frac{q_t^{1+\alpha}}{(1+\alpha)(1+\rho)^t} = \frac{\Lambda^{\frac{1+\alpha}{\alpha}}}{1+\alpha} \sum_{t=0}^{\infty} \left[\frac{1+\rho}{(1+r)^{1+\alpha}}\right]^{\frac{t}{\alpha}}$$

$$= \frac{\Lambda^{\frac{1+\alpha}{\alpha}}}{1+\alpha} \left[1 - \frac{(1+\rho)^{\frac{1}{\alpha}}}{(1+r)^{\frac{1+\alpha}{\alpha}}}\right]^{-1}$$

$$= \frac{\Lambda^{\frac{1+\alpha}{\alpha}}}{1+\alpha} R \quad , \qquad where \qquad (4.13)$$

$$R = \left[1 - \frac{(1+\rho)^{\frac{1}{\alpha}}}{(1+r)^{\frac{1+\alpha}{\alpha}}}\right]^{-1} \qquad (4.14)$$

Equation (4.10), the third FOC above, satisfaction of the lifetime budget constraint and $\lim_{t \to \infty} W_t = 0$ together imply:

$$\sum_{t=0}^{\infty} \frac{q_t}{(1+r)^t} = W_0 + \sum_{t=0}^{\infty} \frac{y}{(1+r)^t} \qquad (4.15)$$

Substituting for q_t from (4.12) into the budget constraint (4.15):

$$\Lambda^{\frac{1}{\alpha}} \sum_{t=0}^{\infty} \left[\frac{1+\rho}{1+r}\right]^{\frac{t}{\alpha}} \frac{1}{(1+r)^t} = W_0 + \sum_{t=0}^{\infty} \frac{y}{(1+r)^t}$$

Summing the infinite series we obtain

$$\Lambda^{\frac{1}{\alpha}} R = W_0 + y \left[\frac{1+r}{r}\right]$$

$$\frac{\Lambda^{\frac{1+\alpha}{\alpha}}}{1+\alpha} R^{1+\alpha} = \frac{\left[W_0 + y\left[\frac{1+r}{r}\right]\right]^{1+\alpha}}{(1+\alpha)} = \sum_{t=0}^{\infty} \frac{q_t^{1+\alpha}}{(1+\alpha)(1+\rho)^t} R^\alpha \qquad (4.16)$$

The first and third terms in (4.16) reflects multiplication of (4.13) by R^α. We can thus use (4.16)

to eliminate Λ, obtaining:

$$v(W) = \frac{\left[W_0 + y\left(\frac{1+r}{r}\right)\right]^{1+\alpha}}{(1+\alpha)R^{\alpha}} \tag{4.17}$$

which is the discounted stream of utility from consumption evaluated at the maximized level of q.

Service levels, s for the depositor encountering s^*

The depositor must spend a period with any bank he visits in order to evaluate whether the service being offered is satisfactory to him. After the initial period with a bank, when service received has been assessed as being at least reservation service level, and the depositor decides to stay with the bank, we may imagine two classes of depositor. A depositor who has maintained an unfavourable record with the bank will expect the bank to judge him unsuitable for better (risky to the bank, convenient for the depositor) future services. For example, a depositor overdrawing without authorization on his current account may not get a cheque guarantee card; one with a cheque guarantee card may not get a credit card. Thus the service quality received at the introductory stage will not be expected to change over time. On the other hand, a depositor who has been a good accountant, has a regular income and has been prudent takes account of the reputation she has acquired with the bank, and expects the service she receives to increase in the future. We denote the expected increase by $\delta > 1$. δ represents improvements in facilities for spatial and intertemporal transactions.[18]

Having found reservation service level s^* [19], the discounted expected value of service is

$$\sum_{i=0}^{i=\infty} \frac{1}{(1+\rho)^i} \frac{\int\limits_{s=s^*}^{s=\infty} sf(s)\,ds}{[1-F(s^*)]} \quad \text{where} \quad \frac{\int\limits_{s=s^*}^{s=\infty} sf(s)\,ds}{[1-F(s^*)]} \text{ is the expected value of service given that it is satisfac-}$$

tory service. If service quality is not expected to increase, the discounted value of service

---

[18] δ could, of course, also reflect the experience the depositor has acquired of the bank's procedures, and hence her ability to extract improved service.

[19] With an infinite time horizon reservation service quality is invariant over time if there is no change in the service he can receive. Thus, once a bank's quality has been rejected the depositor does not return to that bank.

expected by the depositor from staying with the bank is:

$$v^G(s) = \left[\frac{1+\rho}{\rho}\right] s \,, \tag{4.18}$$

where s is whatever level of service is received, provided that it is above reservation level.

For the class of depositors who expect service to improve, whatever level of service of at least reservation level is received, the expected utility of service is given by:

$$v^G(s) = s + \frac{\delta s}{(1+\rho)} + \frac{\delta s}{(1+\rho)^2} + \ldots = \left[\frac{\delta+\rho}{\rho}\right] s \tag{4.19}$$

Consumption and service: returning to $V^G(W,s)$

Since $V^G(W,s) = v(W) + v^G(s)$ where the first term is given by (4.17) and the second by (4.18) when service is time-invariant, and by (4.19) when it increases, the expected value of the discounted stream of utility from staying with a bank is:

$$V^G(W,s) = \frac{\left[W_0 + y\left[\frac{1+r}{r}\right]\right]^{1+\alpha}}{(1+\alpha)R^\alpha} + \left[\frac{1+\rho}{\rho}\right] s \tag{4.20}$$

when service is not improved, and

$$V^G(W,s) = \frac{\left[W_0 + y\left[\frac{1+r}{r}\right]\right]^{1+\alpha}}{(1+\alpha)R^\alpha} + \left[\frac{\delta+\rho}{\rho}\right] s \tag{4.21}$$

when it is.

Finding $V^B(W)$

To find $V^B(W)$ we use the recurrence relation and the fact that, from our assumptions, the value of service expected by the searching customer depends on his knowledge of the distribution.

For the customer searching for satisfactory service the optimized expected value of utility after entering a bank but before observing service is

$$V^B(W) = \max_{q,s} \cdot \left[\frac{q^{1+\alpha}}{1+\alpha} + \bar{s} + \frac{1}{1+\rho}\int_{\underline{s}}^{\infty} V^G(W,s)f(s)ds + \frac{1}{1+\rho}\int_0^{\underline{s}} V^B(W)f(s)ds\right] \tag{4.22}$$

where $\bar{s} = E(s) = \int_{s=0}^{\bar{s}} s f(s)\, ds$ is the average value of service on the market. The discounted life-

time expected utility from searching is this period's utility of consumption and expected utility of

service, plus the discounted expected utility from the next period onwards. The latter quantity is

equal to the discounted utility obtained at a bank if the depositor finds 'good' service there and

stays, plus the discounted optimized utility of leaving the current bank and continuing to search

(i.e. of facing the same problem one period later) times the probability of encountering unsatis-

factory service. Substituting for $V^G(W,s)$ from (4.20) into (4.22), and substituting for W from the

budget constraint (4.2) into $V^B(W)$ on the RHS above, we have:

$$V^B(W) = \max_{q,s} \left[\frac{q^{1+\alpha}}{1+\alpha} + \bar{s} + \frac{1}{1+\rho}\, G(s^*) \frac{\left[W_0 + y\left[\frac{1+r}{r}\right]\right]^{1+\alpha}}{(1+\alpha)R^\alpha} \right.$$

$$\left. + \frac{1}{1+\rho}\, G(s^*)\, v^G(s) + \frac{1}{1+\rho} F(s^*)\, V^B[(W_t + y - q_t)(1+r)] \right] \qquad (4.23)$$

where $F(s^*)$ is the probability of encountering service below the reservation level and

$G(s^*) = \left[1 - F(s^*)\right]$ is the probability of receiving service quality of at least reservation level. In

what follows for notational simplicity, we drop the subscript t on W_t and q_t. Evaluating W_0 at

$(W + y - q)(1 + r)$, (4.23) can be written

$$V^B(W) = \max_{q,s} \left[\frac{q^{1+\alpha}}{1+\alpha} + \bar{s} + \frac{1}{1+\rho}\, G(s^*) \left\{ \frac{[(W+y-q)(1+r) + y\{(1+r)/r\}]^{1+\alpha}}{(1+\alpha)R^\alpha} \right\} \right.$$

$$\left. + \frac{1}{1+\rho}\, G(s^*)\, v^G(s) + \frac{1}{1+\rho} F(s^*)\, V^B[(W_t + y - q_t)(1+r)] \right] \qquad (4.24)$$

Given that V^B and V^G have the same structural form, differing only in the utility of service,

using (4.18) we conjecture that when the depositor has not encountered his reservation service :

$$V^B(W) = v(W) + \left[\frac{1+\rho}{\rho}\right] \sigma \qquad (4.25)$$

where $v(W)$ is the consumption component of $V^G(W,s)$ defined above and σ denotes the level of

service the searching depositor will receive. We may also write $V^B(W)$ as $V^B(W) = v(W) + v^B$,

with v^B representing the second term on the RHS of (4.25). Using this conjecture, we substitute (4.25) for $V^B(W)$ on the RHS of (4.24) and take account of the value of $v(W)$ found in (4.17), while evaluating W_0 as before. Simplifying we have:

$$V^B(W) = \max_{q,s} \left[\frac{q^{1+\alpha}}{1+\alpha} + \bar{s} + \frac{1}{1+\rho} \left\{ \frac{[(W+y-q)(1+r) + y\{(1+r)/r\}]^{1+\alpha}}{(1+\alpha)R^\alpha} \right\} \right.$$

$$\left. + \frac{1}{1+\rho} \left\{ G(s^*)v^G(s) + F(s^*)v^B \right\} \right]$$ (4.26)

As before, it is convenient to analyze the consumption and quality components of $V^B(W)$ separately. We begin by considering the choice of q.

Optimal choice of consumption, q

Differentiating the RHS of (4.26) with respect to q gives the FOC:

$$q^\alpha = \frac{(1+r)}{(1+\rho)R^\alpha} [(W+y-q)(1+r) + y\{(1+r)/r\}]^\alpha$$

Collecting terms in q, using (4.14) and simplifying we obtain the optimizing value of q:

$$q = \frac{1}{R}\left[W + y\left(\frac{1+r}{r} \right) \right]$$ (4.27)

This is the value of consumption derived from pursuing an optimal policy. Thus, R^{-1} can be interpreted as the marginal (and average) propensity to consume out of wealth. From (4.14) we note that this propensity depends on the subjective relative to the market discount factor and the constant intertemporal elasticity of marginal utility of consumption. Substituting for q from (4.27) into the first and third terms on the RHS of (4.26), and simplifying, we find the consumption component of $V^B(W)$, denoted v, to be:

$$v = \frac{\left[W + y\left(\frac{1+r}{r} \right) \right]^{1+\alpha}}{(1+\alpha)R^{1+\alpha}} \left[1 + \frac{(1+r)^{1+\alpha}}{(1+\rho)R^\alpha}(R-1)^{1+\alpha} \right]$$

The complicated-looking term added to unity in the square brackets above reduces to $R - 1$ therefore, as conjectured, compare (4.17), the consumption component of $V^B(W)$, v ,is

$$v = \frac{\left[W + y\left(\frac{1+r}{r} \right) \right]^{1+\alpha}}{(1+\alpha)R^\alpha} = v(W).$$ (4.28)

Finding σ, the service level obtained by the searcher

Having seen that the conjecture regarding consumption in $V^B(W)$ is correct, we find the value of σ in (4.25). Considering s alone, the conjecture is:

$$V^B = \frac{1+\rho}{\rho} \sigma = \max_s \left\{ \bar{s} + \frac{1}{1+\rho} G(s^*) V^G(s) + \frac{1}{1+\rho} F(s^*) \frac{1+\rho}{\rho} \sigma \right\}$$

Substituting for $V^G(s)$ and cancelling $(1+\rho)$:

$$V^B = \max_s \left\{ \bar{s} + \frac{1}{\rho} \int_{s=s^*}^{s=\infty} s f(s)\, ds + \frac{\sigma}{\rho} \int_{s=0}^{s=s^*} f(s)\, ds \right\}$$

σ may be found by differentiating this function with respect to s^*. Alternatively, however, we note that the FOCs for the optimal choice of s^* imply that it is determined where the depositor is just indifferent between staying at his current bank and searching. That is, for the searching depositor, staying with a bank offering reservation service gives the same utility as will waiting to search again, so that $V^B = V^G(s^*)$ where $V^G(s^*)$ is V^G evaluated at s^*. Thus

$$V^B = \frac{1+\rho}{\rho} \sigma = V^G(s^*) = \frac{1+\rho}{\rho} s^* \qquad (4.29)$$

and therefore $\sigma = s^*$: the value of service received by the customer who continues searching is the reservation level of service. Since the searching depositor is moving from bank to bank in his effort to find satisfactory service and his reservation service quality is based on his knowledge of the parameters of the distribution, he will eventually obtain the service quality required.

Consumption and service: returning to $V^B(W)$

From (4.27), (4.28) and (4.29) we may write the total expected value of a searcher's discounted stream of utility from searching as:

$$V^B(W,s) = \frac{\left[W_0 + y \left[\frac{1+r}{r} \right] \right]^{1+\alpha}}{(1+\alpha)R^\alpha} + \frac{1+\rho}{\rho} s^* \qquad (4.30)$$

Equations (4.20) and (4.30)[20] give the expected values of the discounted streams of utility

---

[20] In finding $V(W,s)$, the optimal value of the objective function (or the state function), there has been no need to explicitly consider the second-order conditions because in dynamic programming problems, it has been shown (see Hadley [1964]) that when the objective function (here $q^{1+\alpha}/(1+\alpha)$) is (strictly) concave, so

for customers following their optimal policies of remaining with a bank, and searching for better service quality, respectively. The designation of a reservation service quality effectively divides depositors into two groups: those who judge that the service they have found is at least acceptable and remain with their current bank; and those who are searching for satisfactory service. In the former group those who have handled their bank accounts in a manner satisfactory to their bank will be receiving enhanced service from the second period onwards. By definition, the latter group includes all searching depositors, both those entering the deposit market for the first time, and those who have been in the market, but have not as yet found service at or above their reservation level.

The relationship between financial savings and the rate of interest

The comparative statics of the consumption choice is of interest mainly because it shows that the implied demand for financial savings responds positively to the deposit rate, and that the rate of time preference is the reservation deposit rate, as assumed in the financial liberalization hypotheses. We differ from the liberalization analyses only in explicit examination of the effects of service quality on depositor behaviour. The implied savings supply function has several arguments in common with that of Shaw's money demand (op. cit., p. 61) function[21].

The long-run intertemporal maximum value functions show that for all depositors, the discounted utility of consumption is wealth and the discounted market value of lifetime income

("permanent" income), $\left[W + y\frac{1+r}{r} \right]$, times a factor $\frac{1}{(1+\alpha)} \left[1 - \frac{(1+\rho)^{\frac{1}{\alpha}}}{(1+r)^{\frac{1+\alpha}{\alpha}}} \right]^{\alpha}$, which depends on

the marginal (average) propensity to consume out of wealth and permanent income, R^{-1} (R is defined in (4.14)), and α. We can first show that:

Holding service quality constant, for any individual the reservation deposit rate, below which the consumer will not supply deposits, is the time preference rate.

---

is the state function. Thus the relative maximum of q obtained in the first order conditions is also a global maximum.

[21] We note that McKinnon's demand for money included an investment/income ratio resulting from his capital/money complementarity hypothesis.

To demonstrate this, we note that the optimal value of the objective function, here $V(W,s)$, is a function of the state parameters of the system . In addition, as noted in footnote 21, if the objective function is (strictly) concave, then its optimal value is also concave. Here the objective function is strictly concave in q and hence, $V(W,s)$ is strictly concave in W. This can be demonstrated by differentiating v(W) from (4.17) with respect to W (where the subscripts indicate the first and second partial derivatives of $V(W,s)$ with respect to W):

$$V_W(W,s) = \frac{1}{R^\alpha}\left[W + y\frac{1+r}{r}\right]^\alpha \tag{4.31}$$

$$V_{WW}(W,s) = \frac{\alpha}{R^\alpha}\left[W + y\frac{1+r}{r}\right]^{\alpha-1} < 0 \tag{4.32}$$

since $\alpha < 0$[22]. Thus, $V(W,s)$ is monotone decreasing in W. Now, differentiating either (4.4) or (4.5) with respect to W, noting that $W_t = (W_{t-1} + y - q_t)(1+r)$, we have that when the consumer is following his optimum policy:

$$V_W(W,s) = \frac{1+r}{1+\rho} V_W[(W + y - q)(1+r),s]$$

Therefore:

$$V_W(W_t, s) = \frac{1+r}{1+\rho} V_W(W_{t+1}, s) \tag{4.33}$$

If deposits are made in any period, we must have $W_{t+1} > W_t$. Then, with $V_W(W,s)$ monotone decreasing in W, $V_W(W_t,s) > V_W(W_{t+1},s)$. So for the equality in (4.33) to hold with deposits being made we require $\left[\frac{1+r}{1+\rho}\right] > 1$. That is, the deposit rate of interest must be greater than the time preference rate and the latter is therefore the reservation deposit rate. Given that $r \geq \rho$, so that deposits are being made, we can also show that:

> the value of the sum of discounted utilities of consumption is a positive function of the rate of interest so that the depositor will be attracted by a higher rate of interest, holding service quality constant.

Differentiating v(W) with respect to r and simplifying:

$$\frac{\partial v(W)}{\partial r} = m^\alpha \left[W + y\frac{1+r}{r} \right]^\alpha \left[\frac{1}{z^{-1}-1} \frac{W}{1+r} + \frac{y}{r} \left(\frac{1}{z^{-1}-1} - \frac{1}{r} \right) \right] \qquad (4.34)$$

where $0 < m = \left[1 - \frac{(1+\rho)^{\frac{1}{\alpha}}}{(1+r)^{\frac{1+\alpha}{\alpha}}} \right] < 1$ is the marginal (average) propensity to consume out of per-

manent income and $z = 1 - m < 1$. $\frac{\partial v(W)}{\partial r}$ is unambiguously positive if $\frac{1}{z^{-1}-1} > r^{-1}$. Simplify-

ing, given z's definition, this last inequality implies $(1+\rho)^{\frac{1}{\alpha}} > (1+r)^{\frac{1}{\alpha}}$ and, since $\alpha < 0$, given

$r \geq \rho$, the optimal discounted utility from consumption is increasing in the deposit rate.

A change in the deposit rate has both an income and substitution effect, so that a depositor,

although switching banks in response to a higher rate of interest, may decrease his deposit flow to

the new bank. Using Merton's ([1969], p.254) intertemporal generalization of the Slutsky equa-

tion[23], we can show that:

> the intertemporal compensated substitution effect dominates the income (wealth) effect on
> consumption so that, as assumed in the liberalization literature, total deposits will rise with
> interest rates.

Let $\frac{\partial q}{\partial r}$ represent the total effect on consumption of an increase in the rate of interest. Following

Merton, we may then define $\frac{\partial q}{\partial r}|_{\bar{v}}$ as the Hicks compensated substitution effect. Then, analo-

gous to the usual Slutsky, the intertemporal wealth effect is given by:

$$\left[\frac{\partial q}{\partial r} - \frac{\partial q}{\partial r}|_{\bar{v}} \right]$$

We proceed by finding the two component terms of this expression. Differentiating the expres-

sion for optimal q given in (4.27) with respect to r and simplifying, the total effect on consump-

tion of an increase in the rate of interest is:

$$\frac{\partial q}{\partial r} = -m \frac{y}{r^2} + \frac{1+\alpha}{\alpha} \frac{1-m}{1+r} \left[W + y\frac{1+r}{r} \right] \qquad (4.35)$$

Taking the total derivative of (4.27), holding V constant, and dividing by dr:

$$\frac{\partial q}{\partial r}|_{\bar{v}} = m\left[\frac{\partial W}{\partial r}|_{\bar{v}} + \frac{1+r}{r}\frac{\partial y}{\partial r}|_{\bar{v}}\right] - m\frac{y}{r^2} + \frac{1+\alpha}{\alpha}\frac{1-m}{1+r}\left[W + y\frac{1+r}{r}\right] \tag{4.36}$$

For given V, taking the total derivative of (4.17) and dividing by dr, we find that the first square bracketed term on the RHS of (4.36) is given by :

$$\frac{dW}{dr}|_{\bar{v}} + \frac{1+r}{r}\frac{dy}{dr}|_{\bar{v}} = \frac{y}{r^2} - \frac{\partial m}{\partial r}\frac{1}{m}\frac{\alpha}{1+\alpha}\left[W + y\frac{1+r}{r}\right] \tag{4.37}$$

Substituting from (4.37) into (4.36) and simplifying, the compensated substitution effect of a change in the rate of interest is:

$$\frac{\partial q}{\partial r}|_{\bar{v}} = \frac{1-m}{1+r}\frac{1}{\alpha}\left[W + y\frac{1+r}{r}\right] < 0 \tag{4.38}$$

Subtracting (4.38) from (4.35), the intertemporal wealth effect is:

$$\left[\frac{\partial q}{\partial r} - \frac{\partial q}{\partial r}|_{\bar{v}}\right] = -m\frac{y}{r^2} + \frac{1-m}{1+r}\left[W + y\frac{1+r}{r}\right] \tag{4.39}$$

Thus, the substitution effect dominates the wealth effect if the RHS of (4.38) exceeds the RHS of (4.39) in absolute value. We may safely assume that the intertemporal wealth effect is positive; consumption is a normal good. This implies that the RHS of (4.39) is positive. Thus, if $|\alpha| \le 1$ the substitution effect dominates the wealth effect. $|\alpha| \le 1$, an inelastic or unitary elastic marginal utility of consumption between periods, implies that as consumption increases over time, the rate at which the marginal utility of consumption falls over time increases less than or in proportion to the increase in consumption, as required by the optimality of $V(W,s)$. Therefore, the negative intertemporal substitution effect dominates the wealth effect on consumption of an interest rate change and increases in the deposit rate raises the consumer's deposit flow.

Service quality influences on depositor behaviour[24]: the endogeneization of switching costs

One of our key results is obtained in this section: switching costs are derived as an interest rate premium required to compensate the depositor for the loss of improved service.

We first consider how the reservation service quality changes once the depositor has

---

[24] This section considers service quality at fixed rates of interest until the tradeoff between service and interest is derived.

decided to stop searching and evaluates his expected utility, looking forward one period. Secondly, a utility trade-off between service quality and the rates of interest offered at alternative banks can be derived by approximating the $V(W,s)$ of any individual faced with the choice. We call the interest rate premium required to compensate the depositor for changing banks once he has established a reputation with a bank in long run equilibrium, the switching cost of the individual.

Optimal reservation service, s^*

We proceed by deriving the value of reservation service quality when service quality remains constant, then show how the value of reservation service quality at which $v^G(s^*)$ is evaluated changes when s increases to δs in each period with a bank after the first period.

Considering only the s component of V^B, we have from (4.22):

$$v^B = \max_s \left[\bar{s} + \frac{1}{1+\rho} \int_{s=s^*}^{s=\infty} v^G(s)f(s)ds + \frac{1}{1+\rho}F(s^*)v^B \right] \tag{4.40}$$

Evaluating v^B at its maximum, substituting for $v^G(s)$, solving for v^B and using the same intuitive argument as earlier to obtain v^B (optimization of the expression found for v^B is equivalent):

$$v^B = \frac{\bar{s} + \frac{1}{\rho} \int_{s=s^*}^{s=\infty} sf(s)\,ds}{1 - \frac{1}{1+\rho}F(s^*)} = v^G(s^*) = \frac{1+\rho}{\rho} s^* \tag{4.41}$$

Solving for s^*:

$$s^* = \frac{\rho\bar{s} + \int_{s=s^*}^{s=\infty} sf(s)\,ds}{\rho + \int_{s=s^*}^{s=\infty} f(s)\,ds} \tag{4.42}$$

Recalling that $G(s^*) = 1 - F(s^*)$, the probability of finding service quality above the reservation level, and defining $\psi(s^*) = \dfrac{\int_{s=s^*}^{s=\infty} sf(s)\,ds}{\int_{s=s^*}^{s=\infty} f(s)\,ds}$, the conditional expected value of service, given that it is at least s^*. Then from (4.42)

$$s^* = \left[\frac{\rho}{G(s^*)+\rho} \right] \bar{s} + \left[\frac{G(s^*)}{G(s^*)+\rho} \right] \psi(s^*) \tag{4.43}$$

The reservation service quality is a weighted combination of average quality available at banks and the average good quality, reflecting the depositor's knowledge of the distribution of service quality on the market. The reservation quality level also depends on the depositor's degree of preference for present over future consumption. Taking the total derivative of equation (4.43), or noting that, as $\rho \to 0$, $s^* \to \left[\frac{G(s^*)}{G(s^*)+\rho} \right] \psi(s^*)$ indicates that the optimizing depositor who attaches less preference to current service and is more willing to wait for good service, will tend to have a reservation service quality nearer the level of average good quality. The more patient depositor would be willing to search more and hence reject more market offers, preferring to wait for higher quality.

Consider the optimal choice of s^* of the depositor for whom, after the first period, service quality grows to δs in each period, then substituting from (4.19) for v^G into (4.40), we obtain the expression equivalent to (4.41), v^B for the depositor whose service improves:

$$v^B = \frac{\bar{s} + \frac{\delta+\rho}{\rho(1+\rho)} \int_{s=s^*}^{s=\infty} sf(s)\, ds}{1 - \frac{1}{1+\rho} F(s^*)} = v^G(s^*) = \frac{\delta+\rho}{\rho} s^* \tag{4.44}$$

We may solve for s^* as

$$s^* = \frac{1+\rho}{\delta+\rho} \frac{\rho}{G(s^*)+\rho} \bar{s} + \frac{G(s^*)}{G(s^*)+\rho} \psi(s^*) \tag{4.45}$$

Denoting the s^* in (4.45) by s^{**}, note that $\delta = 1$ implies that s^* and s^{**} are identical. To compare them we take the total derivative of and substitute from (4.45), and simplify to give:

$$\frac{d\,s^{**}}{d\,\delta} = - \frac{(1+\rho)\rho\bar{s}}{(\delta+\rho)^2 (G(s^*)+\rho)} < 0$$

Conditional on having located s^* and an improved service a depositor sets a lower reservation level for service. Since service will improve after his first period with a bank, the depositor's options at the current bank have expanded and the level of service required at the current bank to make him indifferent between searching and staying is lower. But note that this applies only at

the current bank. Even with general knowledge that s increases after a period of reputation establishment, service available at as-yet-unvisited banks remains uncertain. For visits to new banks the long-run discounted expected utility of search after the first period remains $v^B = v^G(s^*) < v^G(\delta s^*)$. Having searched with s^* the enhancement of service means that the set of acceptable qualities is greater because, whatever $s \geq s^*$ has been encountered, the depositor knows that it will improve once he stays with that bank. This implies that, having evaluated s, he is receiving better service than is required to make him indifferent between searching and staying. As a result, for even the marginal depositor the discounted utility of staying with a bank offering known good service exceeds that of continuing to search. As illustrated in Figure 4.1, *ex post* the depositor is in the shaded area of the distribution which was his *ex ante* acceptable set, while the *ex post* set of acceptable qualities is the striped area.

Figure 4.1:
The change in reservation service quality and the set of acceptable service qualities
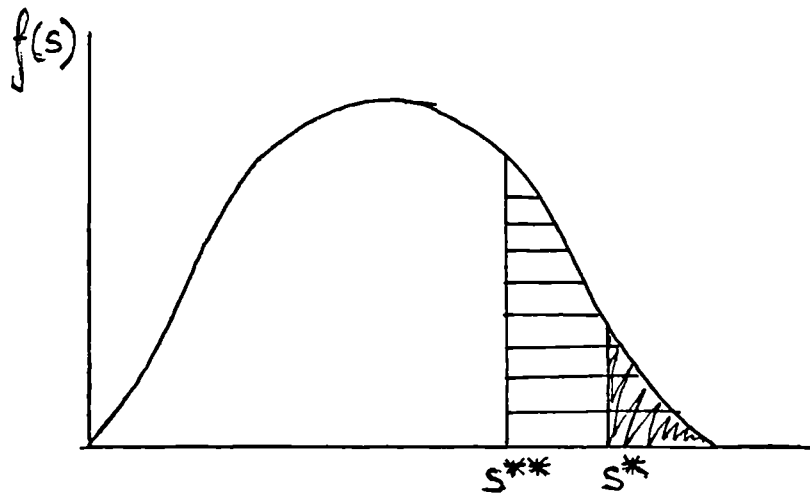
There is an *ex post* matching of depositor and bank which results in the reluctance of even marginal depositors to quit. There is another way of interpreting this which emphasizes the specificity

(matching) of the relationship between customer and bank. In order to observe service and know of its improvement the depositor has had to remain with a bank for one period receiving whatever service is offered, i.e. he has made an investment specific to that bank. To see this note that the cost of another search is the opportunity cost of depositing for one period at reservation service level. With no improvement in service, that opportunity cost is just s^* since the depositor's investment is only in information about s . With improved service the opportunity cost is δs^* since investment in reputation has also taken place. In search models with explicit search costs, a general result is that higher search costs produce higher reservation prices (or lower reservation wages in labour search models) and less search. Here we have an *ex post* increase in the opportunity cost of search and hence a lower reservation service level and less search. The depositor receiving even marginally acceptable service has an incentive to remain with the bank, since he will incur switching costs resulting from this increase in opportunity costs if he changes banks.

Switching costs: the service-interest rate tradeoff

In the short-run in this market there will be searching depositors who have not encountered their reservation service quality. These searchers are easily attracted by higher interest rates since they are unsatisfied with the service received at the bank where they are currently depositing. However, in long-run stationary equilibrium all depositors will have found at least their minimum acceptable service quality since this is determined by the market distribution of service. It is also reasonable to assume that all are receiving improved service since our interest is in depositors' interest rate elasticity of deposit supply: any depositors who have not established a good reputation in the long-run are unlikely to be customers for whom banks would compete. In a mature market the marginal depositor can be viewed as facing two groups of banks: he has already sampled in one pool j where he has found good, improving service, while he is still uncertain of the quality available at banks in pool i . Such a depositor will only be attracted by an increase in the deposit rates offered by banks in pool i if the discounted maximal utility from consumption and service, taking account of the relative financial returns, is higher with a return to search, that is, if

$$V^B(r_i) = V^G(r_i, s^*) \geq V^G(r_j, \delta s^*)$$

This requires

$$\left[W+y\frac{1+r_i}{r_i}\right]^{1+\alpha}(1+\alpha)^{-1}\left[1-\frac{(1+\rho)^{\frac{1}{\alpha}}}{(1+r_i)^{\frac{1+\alpha}{\alpha}}}\right]^{\alpha}-\left[W+y\frac{1+r_j}{r_j}\right]^{1+\alpha}(1+\alpha)^{-1}\left[1-\frac{(1+\rho)^{\frac{1}{\alpha}}}{(1+r_j)^{\frac{1+\alpha}{\alpha}}}\right]^{\alpha}>s^*\left[\frac{\delta-1}{\rho}\right]$$

Since we are considering switches by a single individual the wealth level does not affect this choice and W, y and α are constant. Therefore in order to make this comparison we linearize the expression for $v(W)$ around $r=\rho$, and employing the resulting approximation and simplifying, we find that $V^G(r_i,s^*) \geq V^G(r_j,\delta s^*)$ implies, in terms of the relative interest rates:

$$r_i - r_j \geq \frac{s^*(\delta-1)}{\rho} \frac{\rho^2}{\left[W+y\frac{1+\rho}{\rho}\right]^{\alpha}\left[\frac{\rho}{1+\rho}\right]^{1+\alpha}W} \qquad (4.46)$$

In the long-run a marginal depositor who has sampled bank pool j and found satisfactory service among these will only switch banks if the excess of r_i over r_j is at least as great as the difference $v^G(\delta s^*) - v^G(s^*)$, the gain in discounted service obtained after one period, weighted by a factor which depends on his wealth, permanent income and time preference rate. We call the RHS of (4.46) switching costs, denoted c. With no service improvement $c=0$ for the marginal depositor. Heterogeneous depositors will have different switching costs since these depend on the time preference rate, wealth and permanent income of the depositor. c is obviously decreasing in permanent income and wealth, with the utility gain from the income return to saving being less attractive the lower is wealth. Differentiation of c with respect to ρ gives:

$$\frac{\partial c}{\partial \rho} = \frac{c}{\rho} + \frac{\alpha y}{\left[\left[W+y\frac{1+\rho}{\rho}\right]\left[\frac{\rho}{1+\rho}\right]\right]^{1+\alpha}\frac{W}{\rho^2}} + \frac{1+\alpha}{\left[\left[W+y\frac{1+\rho}{\rho}\right]\left[\frac{\rho}{1+\rho}\right]\right]^{\alpha}\frac{W}{(1+\rho)^2}} > 0 \qquad (4.47)$$

Switching costs are increasing in the time preference rate: the impatient consumer places more value on immediate satisfactory service and requires a higher positive rate of interest differential to compensate him for lost service and induce him to switch banks. Further, as $\rho \to 0$, $c \to 0$: for the infinitely patient depositor who does not discount the future, postponement of the gain that brings about a continuing relationship is costless and individual consumer demand becomes perfectly interest rate elastic.

We may compare this relationship between switching costs and the time preference rate with the buyer response in Gale's [1987] results discussed in Section 4.4. In a mixed price-setting

game, Gale finds that, with a positive probability that the buyer is unattached to the seller (his *ex ante* game), as the buyer's time preference rate tends to zero, price tends to the competitive price. This finding depends on both the increased price elasticity of a patient buyer and the resulting pricing incentives of the seller. Only the buyer response is relevant here. The intuition both here and in Gale's buyer response is simply that the more willing is the customer to postpone consumption (the lower the rate of time preference) the more responsive he will be to price inducement. In the savings context this suggests a paradox which reinforces the switching cost application to the deposit market since (see discussion following (4.33)) it implies that it is precisely the consumer who will accept very low deposit rates who has very low switching costs.

Switching costs' inverse relation to wealth and positive relation to time preference have complementary implications for our application to LDCs. In poor countries the intermediation facilities from the deposit relation may loom larger than the income return in consumer preferences. It may also be conjectured that time preference rates are higher as a result of greater general uncertainty about the future and fewer opportunities for intertemporal trading. Both these factors suggest that switching costs may be higher in LDCs.

We differ from the analyses of financial liberalization in explicit examination of the effects of service quality on depositor behaviour. However, the inclusion of service quality in itself does no violence to, for example, Shaw's hypothesis. Our service quality variable corresponds to the variable reflecting money's intermediation services in his money demand function. Although we assume the depositor must search for his desired services, we would argue that such information acquisition reflects one of the costs of savings disposal discussed by Shaw (op. cit., p.61) if we do not assume, as he seems to have done, that the monetary system was free of such costs and risks.

We have thus endogeneized switching costs from the point of view of the long-run optimal strategy of the depositor. Given our assumptions, in the long-run equilibrium of the banking market there will be no marginal depositors who respond to small interest rate changes. The implications of this result for bank deposit rate setters are taken up in Section 4.5.

Section 4.5: The Banking Market

Banks' deposit rate determination

This section uses a representation of the deposit supply implicit in Section 4.4's model, together with the switching cost variable developed there, to supply the depositor fundamentals for our use in the banking market of Klemperer's [1987a] model of switching costs in a mature market. It concludes that noncooperative deposit-rate-setting banks will find it optimal and feasible to tacitly collude on the deposit rate.

We examine only long-run stationary deposit rate determination. Switching costs do not operate in the short-run - raising the rate of interest will always attract searching depositors, providing banks with a strong incentive to compete; although the attraction or incentive will be neutralized if forward-looking consumers or banks expect to be worse off in the long-run as a result of more elastic demand or more aggressive pricing, respectively, in the short-run (see footnote 28). By focussing on the long-run we are assuming that banks discount short-run depositor switches, looking to the future where there are no shifts because all depositors have found their reservation service and are looking forward to improved service. In the search environment banks recognize this development of switching costs in the long-run equilibrium and, if their discount rates are low, short-run competition will be limited. Furthermore, the depositors who are easily attracted in the short-run are those whose reservation levels are higher than the service found in sampled banks and, as is evident from (4.33) and (4.47), will not only require the lowest deposit rates but will have the lowest switching costs because of their low time preference. The empirical context also supports this treatment: the markets we are considering are mature banking markets with the earliest current banks dating from the late nineteenth century, later entry occurring in the early nineteen fifties and late sixties and most consumers long-habituated to financial institution saving.

Since each bank offers a range of service qualities and depositors do not differ in their preferences among banks, except in so far as they have acquired switching costs, we use Klemperer's [1987a] homogeneous product model to consider banks' long-run profit maximiza-

tion after market shares have been established. Each depositor knows the service at his bank and is looking forward to improved service, and there is no new inflow of depositors[25].

We consider two banks, i and j, who choose their deposit rates r_i and r_j. Take the size of the market as unity and let bank K's share of the market be γ_K, where $\gamma_i + \gamma_j = 1$. Let $P^K(\rho,\bar{c}) = $ the proportion of bank K's ($K = i, j$) depositors with reservation deposit rate, ρ, if no account is taken of switching costs and switching costs, $c \le \bar{c}$. (We assume that depositors are evenly distributed across banks in terms of ρ). Assume $P^K(\rho,\bar{c})$ is differentiable and let $\frac{\partial P^K(\rho,\bar{c})}{\partial c} = p^K(\rho,\bar{c})$.

$\Delta(\rho)$ is the number of depositors with time preference rate ρ, the reservation deposit rate, that is, the market deposit supply function without taking account of switching costs, inferred from the depositor optimization problem. D_K is the deposit supply to bank K. From the examination of depositor behaviour we can conclude that in long-run equilibrium: $P^K(\rho,0) = 0$ that is, all depositors have switching costs and $p^K(\rho,0) = 0$, there are no marginal depositors who are sensitive to small deviations in deposit rates.

Suppose that in long-run equilibrium bank i has a higher deposit rate $r_i \ge r_j$. Then bank i will keep all of its existing customers, $\gamma_i \Delta(r_i)$, and will attract some proportion of j's customers. The latter will consist of: a) those with $\rho \le r_j$ and switching costs $c \le r_i - r_j$; b) those with reservation deposit rate ρ such that $r_j \le \rho < r_i$. They will switch if $\rho \le r_i - c$. Then bank i's deposit supply function in the long-run will be:

$$D_i = \gamma_i \, \Delta(r_i) + \gamma_j \int_{\rho=0}^{r_j} P^j(\rho, r_i - r_j)[d \, \Delta(\rho)] + \gamma_j \int_{\rho=r_j}^{r_i} P^j(\rho, r_i - \rho)[d \, \Delta(\rho)] \qquad (4.48)$$

Since $r_i \ge r_j$, bank j obtains deposits only from its own customers for whom $r_j \ge r_i - c$. Therefore j's deposit supply is:

$$D_j = \gamma_j \int_{\rho=0}^{r_j} \left[1 - P^j(\rho, r_i - r_j) \right] [d \, \Delta(\rho)]$$

$$= \gamma_j \, \Delta(r_j) - \gamma_j \int_{\rho=0}^{r_j} P^j(\rho, r_i - r_j)[d \, \Delta(\rho)] \qquad (4.49)$$

---

[25] Allowing for a very small inflow of depositors need not alter the result if the inflow were too meagre to give any bank the incentive to bid up the deposit rate in order to attract them. Very periodic entry of small, poor households constitutes an example.

Each bank $K = i, j$ maximizes

$$\pi^K = R^K [D(r_K)] - r_K [D(r_K)]$$

where π^K is the profit function and R^K is the net revenue function which depends on K's market share, holding average service quality fixed or assuming that a small change in market share has no effect on the average net revenue from service quality. Deposit-rate setting Nash behaviour implies:

$$\frac{\partial \pi^K}{\partial r_K} = \left[\frac{\partial R^K}{\partial D_K} - r_K \right] \frac{\partial D_K}{\partial r_K} - D(r_K) = 0, \quad K = i, j \tag{4.50}$$

Differentiating D_i from (4.48) and D_j from (4.49) with respect to r_i and r_j, respectively:

$$\frac{\partial D_i}{\partial r_i} = \gamma_i \Delta'(r_i) + \gamma_j \int_{\rho=0}^{r_i} p^j(\rho, r_i - r_j)[d \, \Delta(\rho)] + \gamma_j \int_{\rho=r_j}^{r_i} p^j(\rho, r_i - \rho)[d \, \Delta(\rho)] \tag{4.51}$$

$$\frac{\partial D_j}{\partial r_j} = \gamma_j \Delta'(r_j) + \gamma_i \int_{\rho=0}^{r_i} p^j(\rho, r_i - r_j)[d \, \Delta(\rho)] \tag{4.52}$$

Substituting (4.51) into (4.50) for bank i, the first order conditions are:

$$0 = \left[\frac{\partial R^i}{\partial D_i} - r_i \right] \left\{ \gamma_i \Delta'(r_i) + \gamma_j \int_{\rho=0}^{r_i} p^j(\rho, r_i - r_j)[d \, \Delta(\rho)] + \gamma_j \int_{\rho=r_j}^{r_i} p^j(\rho, r_i - r_j)[d \, \Delta(\rho)] \right\}$$

$$- \gamma_i \Delta(r_i) - \gamma_j \int_{\rho=0}^{r_i} P^j(\rho, r_i - r_j)[d \, \Delta(\rho)] - \gamma_j \int_{\rho=r_j}^{r_i} P^j(\rho, r_i - \rho)[d \, \Delta(\rho)]$$

For bank j's FOCs we substitute (4.52) into (4.50) to obtain:

$$0 = \left[\frac{\partial R^j}{\partial D_j} - r_j \right] \left\{ \gamma_j \Delta'(r_j) + \gamma_j \int_{\rho=0}^{r_i} p^j(\rho, r_i - r_j)[d \, \Delta(\rho)] \right\} - \gamma_j \Delta(r_j) - \gamma_j \int_{\rho=0}^{r_i} P^j(\rho, r_i - r_j)[d \, \Delta(\rho)]$$

In a symmetric equilibrium in pure strategies so that $r_i = r_j = r$, since all depositors have switching costs in the long-run, $P^K(\rho, 0) = 0$, and there are no marginal depositors without switching costs, $p^K(\rho, 0) = 0$, for both banks these FOCs may be written:

$$\left[\frac{\partial R^K}{\partial D_K} - r \right] \gamma_K \Delta'(r) - \gamma_K \Delta(r) = 0$$

If the two banks have constant equal marginal revenues, $R^{i\prime} = R^{j\prime} = MR$ then

$$MR = r \left[1 + \frac{1}{E} \right]$$

where E is deposit supply elasticity, so that the first order conditions for a symmetric (Nash) equilibrium in a market with switching costs are those for a monopsonist, or collusive oligopsony, in a market without switching costs. The same result is obtained if the banks have similar revenue functions and equal market shares. (In the absence of either the latter or constant equal MR, side payments would be necessary in a collusive oligopoly.) This will be a local maximum for each bank if the profit functions are quasi-concave.

The FOC above says that the symmetric Nash equilibrium[26] in a market with switching costs has the appearance of a collusive equilibrium in a market without switching costs, not that it is one. However, it also indicates that tacit collusion of the kind Caribbean economists believe characterizes their banking markets (see Chapter 3) is feasible and sustainable. The literature on collusion points out that it usually fails because firms may have difficulties in agreeing to a pricing rule because their cost functions differ, and lack the incentive to maintain the rule. Enforcement of the rule is also difficult if small deviations cannot be monitored. These difficulties are alleviated by the fact that, with switching costs, the local maximum for each bank acting non-cooperatively is the monopsony solution. With low discount rates, tacit agreement is maintained if each bank simply takes a long-run view of its prospects and serves its own customers, in effect acting as a monopsonist on its market share. None have the incentive to increase deposit rates by a small amount since doing so will not attract more depositors. Only deposit rate changes large enough to cover switching costs are advantageous and these are easily monitored by other banks. Further, with the market broken up into submarkets, a bank losing customers will, especially in the small markets of the Caribbean, be able to tell that the agreement is being broken. We next consider conditions under which large deviations are not profitable.

The collusive outcome as a global maximum

We have found that monopsonistic pricing where each bank sets the deposit rate to exploit

---

[26] Each bank maximizes profit taking the strategy of the other bank as given and prices are equal.

its long-run market share is a local optimum. However, another local maximum may exist which has banks increasing their deposit rate to attract customers with low switching costs. The incentive to do so is obviously greater for banks with a low long-run market share. In the symmetric Nash equilibrium if one bank increases its deposit rate by an amount large enough to cover the switching costs for any proportion of depositors, it not only attracts increased deposits from its own customers but erodes the market share of other banks. Its perceived deposit supply function becomes more elastic at the market (monopsony) price. With banks having the same constant marginal revenue we may then get a distribution of rates with the lowest rate being the monopsony rate. In order to examine the conditions under which the tacit collusion local maximum is also a global maximum we consider the deposit market with two banks paying the monoposony deposit rate, r_m. For simplicity we modify the earlier demand function.

Suppose the market has N consumers with reservation deposit rates, $\rho \leq r_m$ whose individual supply functions are given by $\omega(r)$, so that market demand is given by $N \, \omega(r)$, $\omega'(r) > 0$. There are two banks with market shares $\gamma_1 + \gamma_2 = 1$. Suppose bank 1 contemplates deviating from the tacit agreement to set $r = r_m$. If it raises its deposit rate to $r_v > r_m$ it will attract the depositors with bank 2 for whom $c \leq r_v - r_m$, i.e. all those with switching costs such that these are paid by the excess of its rate over the monopsony rate, as well as maintaining its own market share with increased deposits. Let $\hat{c}$ be the level of switching costs which satisfies this condition with equality. i.e. $\hat{c} = r_v - r_m$, and let $\phi(\hat{c})$ be the proportion of depositors with $c \leq \hat{c}$. (We assume that depositors are evenly distributed by switching costs over the banks). The deviating bank therefore obtains deposit supply

$$D_{1v} = \{\gamma_1 + \phi(\hat{c})\gamma_2\} N \omega(r)$$ (4.53)

while maintaining the agreement he has deposit supply

$$D_{1m} = \gamma_1 N \omega(r)$$

which is just his early share of the market. We note that the elasticity of supply in this latter case is

$$\frac{D'_{1m} \, r}{D_{1m}} = \frac{\omega'(r) r}{\omega(r)} = \eta(r) \, ,$$

where x' denotes the derivative of the function $x(.)$ with respect to its argument; while for the deposit supply in (4.53), the elasticity evaluated at r_m is

$$\frac{D'_{1v}(r_m)r_m}{D_{1v}(r_m)} = \eta(r_m) + \frac{r_m\,\phi'(\hat{c})\,\gamma_2}{[\gamma_1 + \phi(\hat{c})\gamma_2]}$$

$$= \eta(r_m) + e(r_m) \tag{4.54}$$

so that the supply elasticity facing the deviant bank at r_m is equal to the supply elasticity of the individual supply curves, $\eta(r_m)$ plus an additional term, $e(r_m)$. This term reflects the increased market share obtained when he pays the switching costs of those with $c \leq \hat{c}$. Increasing r above r_m thus creates a kink in the supply curve faced by the deviant which becomes more elastic at $r > r_m$ - see Figure 4.2. (Compare Stiglitz's [1987] analysis of a search market with some customers with infinitesimal search costs. Since switching costs here do not tend to zero there is really a discontinuous kink: the increase in elasticity only occurs after the price which pays switching costs and reflects a discrete increase in market share.) Rather than the symmetric equilibrium price which is the collusive deposit rate we may then get a distribution of rates with the lowest the monopsony rate.

Figure 4.2: *Kinked supply curve faced by deviant paying $\hat{c}$ to attract rival's customers*

D_m is the deposit supply curve if the monopsony rate, r_m, is maintained, with MC_m the corresponding marginal expense curve. At deviant deposit rate r_v the supply elasticity increases, with the perceived demand curve being D_v and marginal expense, MC_v
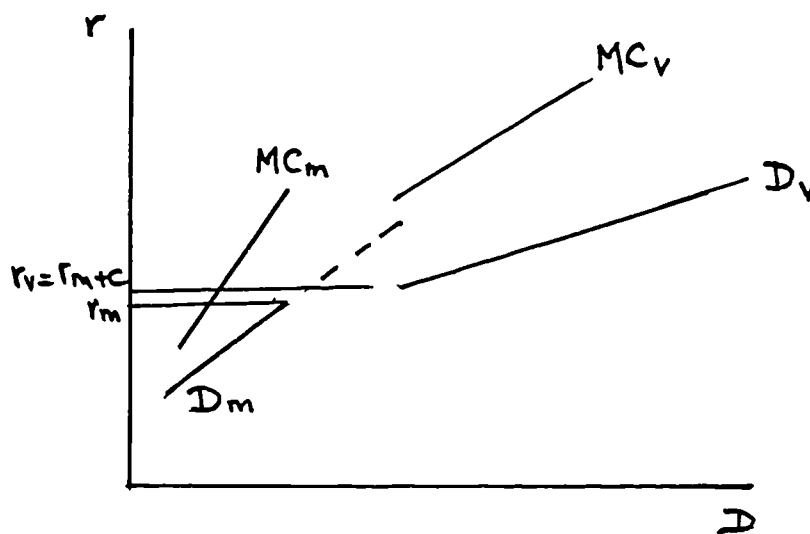
For conditions in which a distribution does not occur, consider the deviant bank's choice of deposit rate premium. Bank 1 chooses $\hat{c}$ to maximize:

$$\pi_{1v} = R\left[\{\gamma_1 + \phi(\hat{c})\gamma_2\} N\omega(r) \right] - (r_m + \hat{c}) \{\gamma_1 + \phi(\hat{c})\gamma_2\} N\omega(r) \tag{4.55}$$

Noting that $\frac{dr_v}{d\hat{c}} = 1$, the FOCs are:

$$R'_v = \hat{c} + r_m + \frac{r_v}{\eta(r_v) + e(r_v)} \tag{4.56}$$

If it were to continue to act as a monopsonist on market share, bank 1's profit function is given by

$\pi_{1m} = R\left[\gamma_1 N\omega(r) \right] - r_m \gamma_1 N\omega(r)$ and the monopsony price is $r_m = \frac{R'_m}{1 + 1/\eta(r_m)}$. Substituting for r_m

into π_{1m} and for r_v from (4.56) into (4.55) we obtain the profit functions which the prospective deviant must compare:

$$\pi_{1v}^* = R_{1v} - \frac{[\eta(r_v) + e(r_v)]R'_v \{\gamma_1 + \phi(\hat{c})\gamma_2\} N\omega(r_v)}{\eta(r_v) + e(r_v) + 1}$$

$$\pi_{1m}^* = R_{1m} - \frac{\eta(r_m)R'_m \gamma_1 N\omega(r_m)}{\eta(r_m) + 1}$$

If $\pi_{1m}^* \geq \pi_{1v}^*$ a bank will not find it worthwhile to deviate from the monopsony agreement. Assuming constant marginal revenue, R', simplifying and substituting for r_m and r_v, we find that the collusive agreement will be adhered to if

$$\frac{\eta(r_v) + e(r_v)}{\eta(r_m)} \geq \frac{(\hat{c} + r_m)\{\gamma_1 + \phi(\hat{c})\gamma_2\} N\omega(r_v)}{r_m \gamma_1 N\omega(r_m)}$$

i.e. if the ratio of increased to simple demand curve elasticities exceeds the ratio of total deviant costs to total costs at the monopsony price. If the increased elasticity is too large the deviant acquires such a high share of the market that it is more worthwhile for him to revert to monopoly pricing.

4.6: Switching Costs and Monopoly Power: Discussion of our results

We have generalized the standard analysis of switching costs by explicitly deriving them from the consumer's intertemporal utility maximization. This endogeneization is significant in

that, by highlighting the difference between switching costs and the transactions costs ubiquitous in economics, it provides an indication of the markets in which the monopoly effects analyzed by Klemperer and others are most likely to bite. In the following we discuss our results in relation to the literature on switching costs and price-setting behaviour.

The monopoly deposit rate outcome conflicts with the usual intuition about price strategies based on the Bertrand[27] model. There the only equilibrium is the perfectly competitive price, even with a small number of firms (in the absence of capacity constraints and product differentiation), because consumers' sensitive response to price differentials induces firms to compete in prices until they are driven down to marginal cost. On the other hand, it recalls Diamond's [1971] result that with imperfect information and positive search costs, the only equilibrium is the monopoly price, even with a large number of firms. With all consumers having positive search costs, each firm can increase its price by ε without losing customers. But once all stores have higher prices, it is feasible for each to raise its price by ε once again. This process results in the monopoly price as the only equilibrium outcome. Some very recent models attempting to reconcile these two stories of price-setting behaviour help to show more generally why the monopoly rate is the equilibrium in the model here. To anticipate, two forces appear necessary to sustain a monopoly price outcome: it must be costly for the buyer to take up lower price offers, and sellers must lack the incentive to compete. If acquiring new price offers is costless, and firms are aware of this and hence have the incentive to compete for sales, we get the Bertrand outcome. If it is costly for buyers to change firm, but sellers still have an incentive to compete, for example, if in the Diamond model there are some marginal buyers with infinitesimal search costs, then the price will tend towards the competitive price.

The key feature of switching costs is that they derive from the loyalty of one agent to another with whom he has already transacted because he anticipates some benefit from the continued association. Transactions costs relate to a cost of trading, over and above the cost of the traded good, whether or not a trade has previously taken place with the current trading partner.

---

[27] The assumption that firms choose price rather than quantity is usually referred to as Bertrand competition after the economist who disputed Cournot's description of quantity setting oligopolists.

While transactions costs are commonly associated with market power, in many cases they seem insufficient to generate a sufficiently strong customer base. As a relevant example cited by Klemperer, there are the transactions cost of opening a new deposit account.

Loyalty is however a backward looking phenomenon and given rational consumers (bygones are bygones) requires some enforcement unless we are to assume that set-up costs have financially constrained the consumer. In our context this is given by imperfect information and future reward. The role of uncertainty in generating switching costs has been recognized by Schmalensee [1982] and Conrad [1983], both of whose consumers are uncertain of the quality of an untried product, and by Klemperer, op. cit., who discusses learning costs. As in Diamond's, op. cit., price search model, imperfect information creates inelastic demand and hence monopoly power. However, search costs are implicitly the opportunity costs of acquiring information and do not have the durability aspect which distinguishes switching costs.

Most models of switching costs' price effects have imposed the costs. In Rosenthal [1982] consumers are non-rational, always purchasing from their original seller unless his price is raised. Klemperer, op. cit., and Farrell and Shapiro [1988] simply assume that a price must deduct the cost in order to attract increased market share. Schmalensee and Conrad both derive explicit utility conditions under which consumers will switch brands. In Schmalensee the consumer familiar with a pioneering brand only buys a new brand of uncertain quality if the discounted expected surplus of buying the new brand (taking account of the possibility that he will have to revert to the old) exceeds the certain capitalized value of continuing to buy his old brand. As in our model knowledge of the new brand is acquired through one period of experience. However, the pioneering brand's advantage in the form of known surplus is given *ex ante*. Similarly, Conrad assumes complete information about the old brand whose advantage derives from the consumer's knowledge. Allowing for *ex ante* uncertainty in both these models would allow (as in a search environment) for marginal customers for whom continued repurchase is not optimal - as recognized by Schmalensee in his Appendix D. With search there are always marginal consumers who provide firms with an incentive to lower price. Transactions costs alone do not create sufficient lock-in to deter rivalry. In the model of Section 4.4, while there are search costs in the form of

the opportunity cost of acquiring information about the availability of reservation service, the costs which make switching non-optimal are the opportunity costs of service enhancement which will be incurred by even marginal consumers.

Klemperer analyzes pricing models with switching costs for both homogeneous [1987a] and horizontally differentiated [1987c] goods. In the first it is necessary to assume that there no marginal consumers sensitive to price deviations in order to obtain the collusive outcome in symmetric noncooperative equilibrium. As this density increases from zero, the profit-maximizing price ranges from the competitive to the monopolistic price in pure strategies. Our demonstration that when service improves after a period of experience even marginal consumers will find it preferable to remain with their old bank avoids the need to impose this assumption and strengthens the monopoly result. In Klemperer's model of horizontal differentiation it need only be assumed that marginal customers are finite and that all consumers have switching costs. Since products are differentiated both by tastes and switching costs, given the limited number of customers attracted by a price cut, the firm does not find it worthwhile to cut price by only a small amount. When tastes change between periods so that the price cut need no longer compensate for both preferences and switching costs, price ranges between the competitive and the monopoly outcome.

In a sense all these models assume an *ex ante*, to the purchase period considered, advantage of the particular brand to which the consumer is attached[28]. An *ex post* advantage explains why even marginal consumers who have just found their reservation level are not easily tempted by the price offers of other firms. It also suggests why even forward looking consumers are willing to commit themselves. Consider contractual or artificial costs discussed by Klemperer. To assume switching costs arising from contracts begs the question of why the consumer would have been willing to sign a legal contract in the first place unless the particular supplier had *ex ante* monopoly power - which prejudges the issue of switching costs' creation of monopoly power.

---

[28] In fact, Klemperer notes that first period competition in a market with switching costs may be fiercer in order to build a customer base for the second period, unless higher market share makes firms more aggressive and forward looking firms wish to reduce rivals' second period aggression, or consumers anticipate the higher prices that will be charged by higher market share firms and respond less readily to lower first period prices.

Our derivation also provides a rationale for the assumption that all consumers in the market have switching costs. In the long-run all consumers with knowledge of the distribution will have found their reservation level. Even if there are consumers for whom service does not increase they are by definition in our context precisely the clients that banks do not consider worth competing for.

Of course, we have ourselves introduced an *ad hoc* assumption in order to endogenize switching costs: it is not shown formally that it is profit-maximizing for banks to increase service quality after a period of learning about the customer. Chapter 5 discusses the justification and provides empirical examples within the banking context.

Switching costs seem more likely to operate where consumers make direct gains from remaining within a particular relationship. Items which are a repeated consumer purchase but provide no additional benefit to their future use - Schmalensee considers a pioneering brand of coffee - seem unlikely to generate such costs. Computer systems are an oft-quoted example of products with lock-in because of the learning involved in their use. They provide an example where we may distinguish between transactions and switching costs. Lock-in arising from learning alone would not be switching costs: some of this learning may be carried over to other systems and if there are some users who have barely acquired the skills they will be attracted by quite low price cuts by others. But consider a user who after search has found the system most suitable to his particular uses and in which packages tailored to him can be developed. There are then future benefits from remaining with that system that can generate switching costs.

Consider Gale's [1987] analysis of monopolistic and competitive tendencies[29]. Gale embeds skeletal versions of the Bertrand and Diamond models in a single mixed stochastic price-setting game. In the " ex ante" version of this game, prices are known before trade and, in both the one-shot and the repeated game, the only equilibrium outcome is the perfectly competitive price. In the "ex post" version, the buyer is attached to the seller before prices are known,

---

[29] The same question has been analyzed by Mortensen [1986] but the conceptual results are similar. There are also several very well-known models analysing the conditions under which there will be price dispersion rather than a single equilibrium price, e.g. Salop and Stiglitz [1977]. Since the main concern here is with the monopoly price outcome, we look only at the models that help explain the latter.

and the unique subgame perfect equilibrium is the monopoly price, even in the case of many sellers. This *ex ante/ ex post* distinction captures the key feature of search costs and switching costs which permits market power: responding to price is costly for the consumer. In the *ex post* model, the consumer must sacrifice current consumption if he is to respond to price, in the *ex ante* model he responds to price freely. In the repeated version of the mixed game in which *ex ante* and *ex post* are played with positive probabilities, (where the buyer knows which game is being played, but the seller does not) price tends to the competitive price as the probability of playing the *ex ante* game increases, and to the monopoly price as the *ex post* game becomes more likely. However, with any positive probability that the *ex ante* game is being played, as the buyer's discount rate tends to zero, price tends towards the competitive price. The buyer with the low discount rate has very low costs of obtaining another price offer, which concentrates the price distribution downwards and, since firms do not know which game is being played, they have the incentive to compete.

As in the switching cost model, poor information and the date of its acquisition are important for market power because they determine the customer's attachment to a particular price-setter. While Gale's one-buyer model cannot capture the heterogeneity of buyers in the market, the case of the low discount rate buyer would correspond to the marginal customer, responding to price differentials. Such customers provide an incentive to firms to compete, an incentive reinforced in Gale's model by the uncertainty of the seller. Switching costs remove this incentive since there no marginal customers in the framework we analyzed (unless there are those whom banks have judged poor prospects).

The effects of marginal customers are also analyzed by Stiglitz [1987] for a duopoly and many-firm industry. He modifies the standard sequential search model by a) allowing for a positive density of consumers with zero search costs and b) explicitly considering the strategic behaviour arising when there are a finite number of firms. Stiglitz's stress on the asymmetry of information arising from the customer's individual knowledge of the price where he is, but not of prices available on the market again points out how information acquisition tends to segment the market. With a continuum of firms, a decrease in price will not bring extra sales, because even

the consumer with small search costs will not find the lower-priced store. With a finite number of stores, however, the small search cost individual will find it worthwhile to search, and price falls below the monopoly price. It is only by taking account of the increased elasticity of market demand, as small search cost individuals are attracted from many stores, that the usual association between competition and many firms is restored. As long, however, as the density of consumers with zero search costs is zero, no search is induced, and the outcome is the monopoly price. It is this outcome that is modelled in the switching cost model above.

Harking back to our earlier description of the requirements for a monopoly price, we may conclude that switching costs permit tacit collusion because they simultaneously make it expensive for the buyer to move for lower prices and remove the seller's incentive to compete. Buyers who have nothing to gain from staying have nothing to contribute to a competitive bank. Search costs and transactions costs, without the matching element known to the price setter, involve a greater degree of seller uncertainty and hence incentive to compete, as in Gale, op. cit.

We have apparently strayed rather far from the repressed financial intermediation context. However, these models show that the association between imperfectly informed customers and monopoly power is a very general phenomenon: whatever the peculiarities of specific models, imperfect observability and hence costly learning (search) reduce consumer price response. Consequently, in the pervasively information-constrained financial markets of LDCs, competitive pricing seems likely to be the exception rather than the rule. Writers on financial intermediation have generally recognized oligopoly power but assume that this will be eroded in the absence of government intervention. We have shown that if noncooperating banks have both the incentive and a feasible environment, the rates they set are most likely to lie below the economy's marginal opportunity cost of funds. In the next section we briefly examine implications for government interest rate policy.

4.7: Policy Implications of a Collusively-determined Deposit Rate

We have considered the effects on depositor and bank behaviour of the improved transactions services a bank offers its clients once they assess each other favourably. The depositor leaves a bank where he has found satisfactory service only if the deposit rate offered elsewhere is high enough to compensate him for the loss of future improved service. This long-run *ex post* matching of bank and client provides banks with stable market shares which, by making monopoly pricing individually profit-maximizing, allows noncooperative banks to operate a tacit collusive agreement.

The monopsonistic deposit rate does not reflect the social opportunity cost of financial savings and reduces the financial flows available for investment financing in countries where potential business borrowers have little or no facilities outside of the banking sector. Thus the market-determined interest rates have results similar to those attributed by McKinnon and Shaw to government usury controls[30].

The orthodox policy correction for the social loss resulting from monopoly pricing is precisely the price controls criticized by financial liberalization theorists. The usual diagram in Figure 4.3 illustrates the effect of a floor on the deposit rate when the collusive duopoly set monopsony rate r_m. L^D is market demand for loanable funds, downward-sloping despite the constant net marginal revenue (NMR) of the two banks because we assume that the revenue function reflects the effects of service quality, administrative expenses etc. With r_F, the deposit rate floor, the effective supply curve facing a bank becomes elastic up to its intersection with D'. The supply of deposits increases and there is less rationing of investment funds.

---

[30] Which is not to say that loan rate ceilings are not also responsible for such outcomes. But they are not alone. And a well-founded observation of collusive behaviour may explain why monetary authorities have a penchant for interest rate controls.

Figure 4.3: *A minimum deposit rate policy which increases financial savings.*
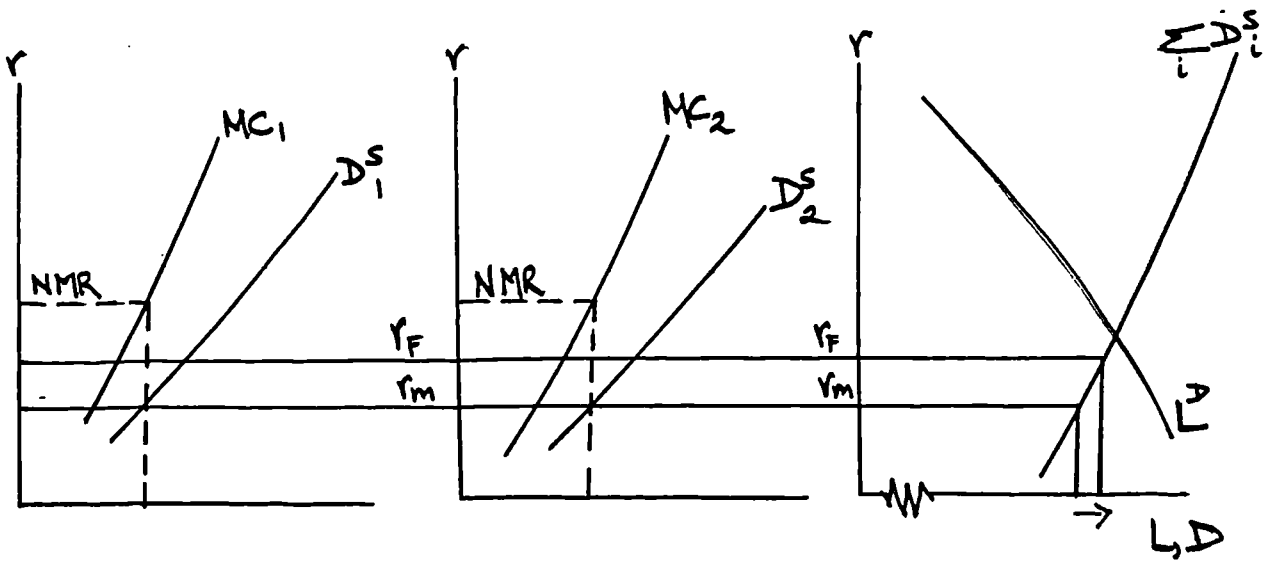
But, as is equally well-known, choice of the minimum rate is unlikely to be easy. Too high a floor may create excess supply and/or induce the bank's exit. The latter seems improbable given the usually close communication between banks and monetary authority in LDCs. Barbados, the only one of our four Caribbean countries to practise systematic interest rate controls, introduced a floor for the first time in 1978. It has been adjusted at intervals in line with international rates and interbank consultations.

Financial liberalization theorists have criticized not deposit rate floors but loan rate ceilings which are far more prevalent. Ceilings are just the counterpart to a floor in a monopolistic loan market and if monopsony power in the the deposit market tends to create monopoly power in the loan market (see later discussion of Yanelle [1986]), a loan rate ceiling may be justified. But fixing the loan rate interferes radically with banks' ability to fix risk-appropriate rates and screening contracts (see Chapter 7). The resulting inefficiencies in loan allocation may outweigh any gain from expanded loanable funds. And, as Chapter 8 shows, taxes are the appropriate instrument for correcting the externalities generated by the screening contracts themselves.

There is another caveat to the apparent endorsement of interest rate controls. Recall that the FOCs for the switching cost model were those of a collusive oligopoly on the assumption of equal market shares or constant equal marginal revenues (MRs). In the absence of such constant MRs a collusive oligopoly with unequal market shares would have to agree on side-payments to members with lower marginal revenues. Official fixing of a loan rate would go a long way towards the equalization and constancy of MRs (if they were not already so) greatly facilitating collusive deposit rate pricing.

The last two paragraphs suggest that efficient policy requires an explicit formal model of the strategic interconnections between the loan and deposit markets. Industrial organization's focus on product markets has diverted attention from markets where firms behave strategically on both input and output determination. As far as I am aware, only Yanelle's, op.cit., paper addresses the question. Very generally, her analysis suggests that such markets may have more scope for the creation of monopoly power than one would expect without intermediation because an element of domination on, say, the input side, allows the intermediary to dominate the output side as well. This may help explain why we tend to observe oligopolistic financial markets[31].

In summary, the collusive behaviour hypothesized in LDC financial markets is derivable from rational consumer behaviour and noncooperative banks' profit maximization in an uncertain world. As Section 4.4 shows, the deposit function we have employed does not differ from that used by financial liberalization advocates. Therefore the results obtained follow from their own priors, viz. interest sensitive and service responsive deposits, imperfect information, the inter-mediation functions of banks. We have merely formulated the results of banks' intermediation services at the individual agent's level, applied a search framework (to take account of one aspect of imperfect information), and considered strategic behaviour (to take account of rational bank profit-maximizing). And, contrary to the usual advice, interest rate regulation can be justified in such markets although several caveats are in order. Rational policy in this context requires careful examination of the markets concerned and, in particular, does not suggest that simple

---

[31] Note also that Yanelle's model did not incorporate the informational constraints that are fundamental to financial markets and which, as we have seen, generally assist in price exploitation.

deregulation will bring the investment expansion and growth usually claimed. Chapter 5 examines the service quality question from banks' perspective and Chapter 6 goes on to enquire why entry does not help to erode market power.

Chapter 5: Implicitly Contracted Service Quality

5.1: Introduction

We have shown that depositors will develop switching costs in the long-run equilibrium of a search market if changing banks means that they lose an improvement in service at their current bank. The object of this chapter is to describe why banks have an incentive to increase service quality. Our argument is as follows. There are benefits to a long-term association between bank and client because the bank can provide transactions services, particularly intertemporal, and insurance services which customers value. Realization of their benefits requires that the services be appropriately tailored to the preferences and means of the client. However, customers have private information about themselves and the value they attach to the bank relation. We may regard bank and client as signing an implicit contract at the start of the deposit relation. The contract allows for the credible transmission of the information from client to bank by specifying a performance-contingent level of service. Service rises over time because higher contingent service reflects data processed about the client. The additional element modelled in Chapter 4 was that rising service quality creates customer loyalty. Loyalty or switching costs can then be viewed as the means by which the bank retains customers with profit potential. The contract is in effect self-enforcing.

It has been usual to take the view that a long-term contract remedies the exploitation made possible when a trader has switching costs which are given *a priori*. Here we argue that it may be fruitful to take the converse view that the creation of switching costs is an endogenous means of enforcing a long-term contract required for repetitious transactions where it is difficult for an outside party to judge the appropriate structure of trades[1]. This view is suggested by consideration of service quality as a variable analogous to those in models where time-varying contract variables serve as a means of contract enforcement. Our view also reflects the bank-customer relation analyzed in early credit rationing literature. In that literature the bank-customer relation is

---

[1] For example, while not inconceivable, it would be very costly for courts to judge the appropriate levels of account performance and service provision between bank and customer.

usually assumed to exist as a result of bank knowledge of customers or because of jointness in the demand for bank facilities. However, while the banking firm literature describes why the bank values a continuing relationship, most models do not explain how it achieves it.

Section 5.2 describes what we mean by service quality and discusses the bank-customer relation. Some authors (for example, Fried and Howitt [1980], Flannery [1982]) have remarked the parallels between the notion of a long-term relation in the bank market and the contracts usually analyzed in the labour market. However, with those partial exceptions, the banking literature does not appear to have exploited the similarity. Section 5.3 discusses models of long-term contracts, arguing that the benefits which these models derive for long-term contracts have parallels in the banking market which help explain banks' motivation to provide increased service quality. Section 5.4 gives a simple informal example of banks' incentive to increase service quality when customer type has been learnt. Section 5.5 discusses self-enforcement and our view of how switching costs achieve this.

5.2: Service quality enhancement in the banking market

We are arguing that the necessarily long-term relationship between bank and customer is governed by an implicit long-term contract which starts on the deposit side. This contract specifies certain behaviour by the client and a contingent enhancement of service quality by the bank. It complements the legal contract which determines price and repayment (deposit and loan) by providing for the quality aspect of the trade. The contract is self-enforcing in that the effect of its service quality specification is to create switching costs for both depositor and bank. This notion of a repeated relation between bank and customer is widespread in the banking literature. We discuss the most closely related[2] instances of it below before describing our idea of service quality and its enhancement.

The idea of better service because of a durable bank-customer relation has mainly been used to explain credit rationing. Fried and Howitt [1980] argue that there is an implicit contract whose

---

[2] Chapter 4 also refers to these ideas.

terms specify that risk neutral banks protect their risk averse clients from loan rate fluctuations and are compensated by borrowers' willingness to pay a higher loan rate[3]. The contract is enforceable because switching banks is costly for bank and customer. They identify the source of these switching costs as bank savings on administrative and screening costs in lending to old customers. This assumes that the customers have remained in place. Except in so far as old customers are less likely to be rationed, clients' incentive to remain at one bank is not analyzed. In addition, the idea that the deposit relation starts on the deposit side is only implicit. On the other hand, Flannery [1982] considers only the deposit side, viewing retail deposits as a factor of production in which the bank has invested in a manner similar to the training Becker [1962] (see Section 5.3) saw firms as providing. As a result, banks pay retail deposit rates above the marginal cost of deposits in earlier periods in order to retain deposit accounts in which they have invested. This analysis stresses the value of the deposit relationship but accounts only for retail deposits and does not consider if or how that relation affects the customer as borrower. Devinney [1986] considers only the loan market, describing the customer relation as the continued willingness of bank and borrower to sign loan contracts with each other. Repetition improves banks' inference about borrower quality and competition induces all banks to offer better contracts to non-defaulters who therefore have no incentive to change banks. Although this characterization of the bank-customer relation is based on the informational advantages which we stress, it does not allow for the feedback from the depositor side of the relation which is generally seen as an important component[4]. It thus ignores what would appear to be an important learning opportunity for the bank, and one which could be cost-saving - learning only by the observation of default seems a costly option. Nor is there any relationship specific advantage of the durability since default is observed on the market; the preferential treatment of established customers, which observers of the bank market consider important (see Blackwell and Santomero [1982]) is lost.

By focussing on service quality enhancement over time in markets where depositors search for service, we believe that we capture observed features of the relationship while explaining why

---

[3] This gives rise to credit rationing because the response of the loan rate to market fluctuations is reduced.

[4] See Wood [1974], discussed in Chapter 4, and Flannery, op.cit.

it is sustainable.

Service Quality

By bank service quality is meant the dimension of the total bank product which permits banks to improve (or not) the agent's ability to transact over time and space. Hence better service quality is an increase in an agent's ability to relax his periodic budget constraint and rearrange (smooth) his consumption pattern over time, as well as to reduce transactions costs in any period. Most generally, banks provide depositors with access to liquid assets that they might otherwise have to hold in the illiquid, risky form of loans (or real goods). Where the liquid alternative may be cash, which is especially relevant in LDCs, banks even provide accounting and security services: savings book entries (for example) help keep track of receipts and expenditures. Different clients require different methods in order that transaction ability is improved, an obvious distinction being that between firm and consumer. This description begins to resemble a horizontal notion of differentiation. Note, however, that a bank provides a specialized service of transaction facilitation and asset transformation. It has knowledge of how a particular client's preferences can be satisfied, given his constraints, that the client himself may not have. The correct tailoring of service thus represents a dimension of a deposit account which would be unanimously ranked by depositors and some banks may be better at such tailoring than others so that their deposit accounts would, at similar prices, be preferred by all depositors. This tailoring requires that the bank acquire information about the customer which some customers may have an incentive to misrepresent.

Improving Service Quality

In order to illustrate how service quality can be increased, three major areas are considered. These are payment methods, credit and international payments. Take payment methods. Credit cards and cheque guarantee cards provide their users with a facility to arrange purchases more cheaply and conveniently over space and time, and are extended on the basis of criteria similar to those used for loans, although at a rather lower level. They also represent a lower service quality since they provide very much less scope for intertemporal allocation than does a loan (indeed, a

cheque guarantee card provides none). For those with demand deposit accounts (passbook savings accounts have tended to dominate in LDCs), a cheque guarantee card extends the uses of a demand deposit account and hence the service quality it embodies. If a credit card is used for credit, a bank earns revenue on its use; on the other hand, losses are possible if the card is unwisely issued.

Consider credit next. Observation of how a customer handles payments methods will assist the bank in decisions about whether or not to extend credit and a positive decision increases quality. But the type and conditions of credit are subject to wide variation. The greater the speed with which a bank is willing to extend a loan, and the more flexible the loan arrangement, the higher the quality of service. Not only will the agent be able to effect his transactions more rapidly but he will also be able to save on information production for the bank. There are also direct savings depending on the type of credit facility offered. For small business, a loan account which commits the client to paying interest on the amount lent on the account, rather than paying as he draws down, as would happen with an overdraft, is obviously far less flexible. Some banks also have facilities which cater to particular groups, with specialized loan officers who are able to advise and offer special assistance to clients. Small business and agriculture are two of the categories commonly available. Whether and what type of security is required for a loan will also vary - with highly assessed customers having greater access to loans without security or with less restrictive forms of security. A further example is provided by the case where a group of companies are allowed to set off credits in one company account against debits in another. This dissuades the company from holding accounts at other banks, at the cost to the subject bank of foregone interest on overdrawn accounts.

Business customers' transactions are also facilitated through the issuance of, for example, letters of credit. A letter of credit represents an improvement in payments facilities because it reduces the risks of international trade between companies (importer and exporter) who are not well-acquainted with each other. Such an improvement in quality will not however, be provided by a bank to an importer unless the bank is satisfied of his client's "creditworthiness".

5.3: Contract model explanations of tenure and its rewards

We describe the benefits attributed to long-term contracts and discuss how these apply to the banking. The idea that longevity or loyalty in economic relations is both valuable and gives rise to fixed costs which produce monopoly distortion has been around at least since Lewis [1949]. Its sources in terms of human capital investment were first explored by Becker [1962], whose insights have informed several recent analyses, and its effects have been studied by Rosenthal [1982], von Weizsacker [1984], Klemperer ([1986a] etc.) and Farrell and Shapiro [1988]. The analysis of how long-term relationships are implemented and of their benefits is provided in numerous models which compare long-term with short-term contracts. Many of these are described in Hart and Holmstrom [1987]. Contract models have the advantage of deriving loyalty, in the sense that they demonstrate the advantages of long-term contracts. However, to the extent that the commitment necessary for the contract is assumed, loyalty is not assured, except when third party enforcement is feasible. If not, the contract itself must provide some surplus which the partners lose if they renege. Their main application has been in the labour market where observed tenure between worker and employer is explained principally by asymmetries of information between worker and employer, on-the-job training or worker insurance.

In the general formulation of the long-term contract model[5], the agent has private information about some random variable (the state of the world), or undertakes a private action, which influences his utility. He chooses a state-dependent action (message) observed, or the results of which are observed, by the principal. The agent chooses a contract among those offered by the principal. Roberts [1982] distinguishes three cases: where the probability of a state in any period is dependent on past state realizations (serially correlated states), where the probability of the state depends on past actions (behaviour dependent states) and where that probability is independent of both past states and actions (time-independent economy). Finite long-term contracts (when commitment is possible) are shown to improve upon short-term contracts in all cases because later dates allow the spreading of risk. We suggest that there are similar reasons for

---

[5] By far the majority of these are principal-agent models, assuming private information. Exceptions will be noted.

long-term contracts to be beneficial to a bank. Banks face risks in most client transactions. This is obvious with credit but even deposits are risky in the sense that early withdrawal and fluctuations make decision-making more difficult. This may partially account for bankers' concern with 'core' deposits (see Flannery [1982] who attributes the concern to their specific investment).

With serially correlated states (for example, where there is hidden information about type) the principal can learn from observation of first period actions. As a result, he will be able to extract an agent's surplus in future short-term contracts. The agent's realization of this distorts his current actions (the ratchet effect). Long-term contracts which commit the principal to future reward can remove such inefficiencies. A similar result is obtained in Baron and Besanko [1984] while Laffont and Tirole [1988] study similar structures when only short-term contracts are available[6]. In the banking market information is also acquired in the early stages of the account relation. The difference arises because both bank and customer gain from the future reward so that a long-term contract need not require explicit enforcement. Depositors differ by "type": in their honesty, management ability, wealth and income, *inter alia*. Type is private information to depositors for some of whom misrepresentation may seem advantageous, while the profitability of the bank's revenue earning services depend on the type to which they are extended. In the opening stage of the deposit account relation, the service extended is constrained by the bank's lack of information; as the bank learns more about the depositor it is able to condition service on that information, adjusting both the components and their proportions to customer needs and capabilities.

Where states are behaviour dependent in such a way that the principal's expected utility is increased by certain past actions, long-term contracts which shift some of the gain to the agent in preferred states improve on short-term contracts. With the latter, the agent's knowledge that the principal's second period payment to the agent (in a two period framework) will reflect his first period action causes the agent to change his first period action so as to influence the second period payment. Long-term contracts that commit the principal to second period payments which

---

[6] Both of these analyses are conducted in a regulator- as-principal, firm-as-agent framework.

reward first period actions desired by the principal induce the agent to take the desired first period action. Baron and Besanko also consider the moral hazard problem when future states are dependent on past actions by the agent, and find that contracts which commit the regulator to decisions which increase the future marginal return to current actions are desirable. The degree of current improvement rises if current actions are observable (no moral hazard). (See also Lambert [1983]).

In the banking market an increase in later service rewards current behaviour. A bank has some standard of a desirable account stemming from both its direct contribution to expected profits and the informational contribution of its status. Large, stable deposits contribute directly to reduction in liquidity risk and may reduce administrative costs. Deposits of a given size by a single customer have lower relative costs than deposits of similar aggregate size by several customers. The latter customers may also be transacting more through other institutions which reduces the information base for the bank. The more transactions a customer channels through the bank the greater the informational content of the account. For example, the deposit of salaries provide a direct source of information. These transactions can also of course increase non-interest earnings. If there is a future reward for current activity, the customer's future utility will depend on current activity and more care will be taken with account transactions at the margin. This improves the probability of non-overdrawn accounts, lower administrative costs etc. For instance, the customer who knows that he may wish to borrow from the bank later is less likely to misuse his cheque guarantee card. Again, the contract enforcement requirement is reduced because having monitored the account, the bank has acquired information which will improve its future decisions and hence profits. If the depositor leaves at the end of the monitoring, this advantage is lost. Positively related quality and tenure enables the bank to gain from its information.

Some of the advantages mentioned above for the banking market would also apply if expected future profits were independent of current action - for example, better current care reduces the probability of high current administrative costs for banks. In a time-independent economy, if the principal gains from the contract, the long-term contract improves on the short-term by rewarding present action (effort) by future gain, especially where monitoring is difficult.

This general result provides a rationale for positive tenure-earnings and effort-promotion schemes.

Several labour market models derive similar time-varying contract variables though the sources of the duration-linked payment differ. In Becker [1962] higher wages reflected productive training or, in the case of specific training, encouraged worker participation and discouraged employer layoffs. Salop and Salop [1976] have high future wages as a screening device to induce more productive workers to self-select; Guasch and Weiss [1980] interpret a lower first period wage as workers' payment of a testing fee which discourages applicants who know themselves unlikely to successfully complete the apprenticeship. Implicit contracts (enforced by reputation) which specify wages first below and then above the value of marginal product are also seen as discouraging shirking (Lazear [1981]). Holmstrom's [1983] model, however, explained below-marginal-product first period wages as a premium paid by mobile risk averse labour for insurance against income fluctuations (layoffs). The implicit contract assumed in his model is enforced by the firm's need to maintain a reputation for reliability.

There are thus several reasons for tenure rewarding long-term contracts. We have tried to indicate how these may account for benefits to banks. Banks are the consumption smoothers par excellence and as such require a repeated relationship with their customers. But the relationship is fraught with asymmetries of information. A long-term contract helps resolve the asymmetries by exploiting the extra dimensions provided by memory in a necessarily multi-period relation. We have also suggested that since the bank reward, does not, like wages, represent solely a cost to the bank, customers may be confident that the promised reward is forthcoming.

5.4: An Example of improving service quality

An example shows how a bank stands to gain if it extends improved service to customers after it acquires information on their requirements and constraints. Suppose a subset of depositors consisting of two *ex ante* indistinguishable groups. Both groups require loans and other revenue-generating services in order to undertake a project. Members of one group have skills in

project management, accounting procedures etc. or are simply more conscientious and careful managers. Members of the second lack these skills or are more cavalier managers[7]. Both request the same loan. Although both groups are likely to claim credit management skills in order to acquire a loan, the bank cannot trust such claims without verification. Members of the second group can successfully be extended loans if the bank provides them with special monitoring and assistance as well as the loans. On the other hand, if the bank takes no special care, their projects will be unsuccessful and the bank incurs losses.

We denote the second group as the NACs (needing assistance and care) and suppose that there are a proportion γ of this group. There are $(1-\gamma)$ of the first group, CAN (who *can* take care of projects without special assistance). When a loan and other services are extended to the CANs, the bank obtains a return r. If a similar facility is extended to the NACs, and the bank also provides special assistance, gross bank return is $r^* > r$. The return is higher because the NACs will require additional revenue-generating services which their limited knowledge prevents them from foreseeing[8]. On the other hand, if NACs are not assisted by the bank's specially trained loan officers, their project fails and the bank loses, i.e. its return is $-R$. Let C be the cost to the bank of offering special assistance and monitoring, where $r < C < r^*$

Suppose a bank were to extend loans and services to recently acquired depositors about whom it has not yet acquired any information. Knowing that the undifferentiated subset of depositors is made up of the two types, it has two options. It can provide special assistance to all clients, earning a total return of;

$$\gamma r^* + (1-\gamma)r - C = \gamma(r^* - C) + (1-\gamma)(r - C) \leq 0$$

Alternatively, rather than incurring the cost of special assistance, it can take its chances with the two groups. In this case it earns:

$$(1-\gamma)r - \gamma R \leq 0$$

---

[7] There may be alternative interpretations in terms of the 'technology' of funds management within a particular business context, so that it is not a question of vertically, but of horizontally, differentiated types.

[8] We can, for example, imagine the NACs to be non-traditional producers of goods which have an export potential and whose future activities may therefore require letters of credit, foreign exchange transactions etc.

The inequalities indicate a supposition that neither of these options are profitable. In that case, the bank would not wish to provide loans before identifying the groups. If it signs an implicit contract to provide improved service which depends on the account performance observed during the depositor-only relationship, it will learn customer characteristics, only offering special facilities to those in need of them. Bank returns are in this case are:

From the NAC group: $r^* - C > 0$

From the CAN group: $r > 0$

Thus no loans are made in the first period but an appropriate credit package is offered in the second, providing improved service. The bank is able to tailor the loan facility to client characteristics. The information acquired through experience with the client thus allows the bank to make profit-maximizing decisions. It therefore has an incentive to improve later service and the promise of enhanced service is credible.

We may also ask why the client does not open deposit accounts at all banks and hence enable all banks to offer improved service, thus avoiding the development of switching costs. This seems unlikely for three reasons. Within the context of this example, first, the client's incentive to do so is reduced by available loan uses. If the project undertaken is replicable with no diseconomies of scale, or the borrower has alternative profitable opportunities to employ the funds, obtaining several loans will be attractive. It seems overly optimistic to assume that borrowers have such a range of profitable opportunities available. If not, borrowing more than required is unprofitable, unless default is intended. Second, from the point of view of the bank, a component of its information gathering about a client is his debtor position. Clients discovered to be in the process of acquiring multiple loans are unlikely to be well-regarded because, as noted above, this increases likelihood of default. There is a third aspect. In order for the bank to learn about a client and to judge the information acquired as reliable and reasonably complete, the deposit account size and transactions volume would have to be at or above some minimum level. If too low, the bank may well refuse to extend improved service on the grounds that it has too little reliable information; or extend very little improvement (credit card but not a loan). We may

speculate that cases of low information content correspond to bankers' arguments that some clients have not shown themselves to be creditworthy or to have bankeable projects.

5.5: Self-enforcement

In order for an implicit contract to be self-enforcing, it must be individually rational for the participants to agree to the contract signing and their continued participation must remain individually rational over the life of the contract. The association must generate a surplus over and above what the participants could obtain by termination (see Macleod and Malcomson [1989]). Reputation is the source of surplus often considered (see Bull [1987], Holmstrom [1983]). Thomas and Worrall [1988] show how time varying payments can enforce implicit contracts required to govern long-term relationships where precommitment is impossible. The incentives to renege on the contract when spot market payments are favourable are alleviated by varying the contract wage[9] so that it remains within a band of the current spot market wage. This results in serially correlated wages. They also show that later rewards (the Holmstrom, op.cit., result), that is, rising wages, occur when the agent who gains little from the contract must be induced to stay on. This result provides some support for our argument when we consider the depositor option to change banks.

The deposit contract normally considered, which is legally enforceable, specifies that the borrower (bank) will pay on demand (or at a specified date) the deposit made by the lender (depositor), together with the accumulated market interest or the interest determined at the time of deposit. But the quality of the account is not enforceable by a third party. It depends on information (monitoring results) and assessments that a third party does not have. That is, the legal deposit contract is incomplete. Another implicit long-term contract would help if it is self-enforcing. The example above showed circumstances under which the bank would have an incentive to increase service. Chapter 4 showed that the searching depositor would have an incentive to remain in the contract with enhanced service. The switching costs make trade with a third party

---

[9] Thomas and Worrall consider the contract required when a risk-averse worker is insured against wage fluctuations by risk neutral firms.

disadvantageous and the very reward to performance which creates switching costs makes adherence to certain performance (no unauthorized overdrawing, up-to-date payments on credit cards etc.) individually rational. But, as shown in Chapter 4, they also facilitate collusive behaviour.

A different view of long-term contracts in a relationship with switching costs is therefore taken by Farrell and Shapiro [1989] [10] who are concerned with the opportunism possible when buyers are locked into sellers because of relationship specific investment. They argue that long-term contracts may remedy such opportunism by pre-specifying price but show that an incomplete contract, specifying the verifiable price, is only an improvement on the arrangement without a contract if utilities are separable in the variables of interest. Otherwise exploitation in the uncontrolled variable will eliminate the gain obtained from the contracted variable. Recognition that a deposit provides a depositor with service facilities in addition to a direct monetary return suggests that the existence of the enforceable deposit contract, which specifies only price and term, is usefully supplemented. [11]

The forward-looking consumer's willingness to sign a contract which will subject him to later price exploitation could be explained by a marginal expected tradeoff. There may, however, be another explanation. The contracts Farrell and Shapiro consider are bilateral contracts but the price exploitation examined in switching cost models is a market phenomenon. It seems likely that a bilateral contract will specify that trade takes place at the market price, with some adjustments subject to choice which satisfy the incentive compatibility constraints. While each individual (depositor) signing the contract may well foresee the opportunities it provides for tacit collusion, if he is small relative to the market he will judge that his refusal to sign the contract will have no effect on market price but will harm him. Competitive prices have a public good aspect. This makes the idea of exploitation-remedying long-term contracts less convincing. Farrell and

---

[10] Our view is closer to that of Aghion and Bolton [1987] where contracts between buyer and seller both lock them in and enable them to extract an entrant's surplus.

[11] Farrell and Shapiro also show that if switching costs are unobservable, incomplete price long-term contracts may improve on a short-term contract by making it so inefficient for the seller to distort quality in order to compensate for the contracted price that he may even refrain from doing so. Where switching costs are endogenously determined precisely because of private information, the non-observability of the costs themselves is not appealing.

Shapiro do not take account of the possibility that long-term contracts may themselves create rather than be a corrective for switching costs. Similarly, Hart and Holmstrom, op.cit., argue that lock-in can motivate (rather than maintain) long-term contracts. For example, Grout [1984] shows that there may be inadequate relationship specific investment in the absence of a long-term contract because the investor cannot appropriate all gains. We are arguing that, in some contexts, it is the relationship-specific investment itself (in information by the bank who must wait to offer profitable services, and in both information and reputation by the customer - see Chapter 4.4) which permits the profitable and credible service improvement that in turn enforces the long-term contract.

5.6: Concluding Remarks

The association between bank and client is by its very nature long-term: borrowing and lending (and deposits are loans from the depositor to his bank) are intrinsically intertemporal. At the most basic level this is because the long-lived individual is concerned with consumption smoothing over his lifetime. At the same time the transactions are fundamentally characterized by asymmetries of information. The implications of this have been most widely analyzed in the credit rationing literature (see Chapter 7). The deposit contract has also been explicitly examined in the Diamond-Dybvig- [1983] inspired models (see Chapter 2). Both contracts commonly and simultaneously govern the association between a single agent and the financial intermediary: depositors borrow from the banks in which they have deposited, the insured borrow on their life insurance policies. The argument here is that the two facets of the association complement each other because of the asymmetry of information that characterizes capital markets. Their functioning requires that information be credibly transmitted. The long-term depositor relationship facilitates this but in order for the relationship to be sustainable it must be enforceable: time-varying service simultaneously incorporates information gained and provides an enforcement mechanism by rewarding continuity. It is precisely because the deposit account provides information that service can improve and that improvement ensures the customer loyalty which permits the bank to benefit from the information.

Chapter 6: Entry and Exit with Imperfect Information and Sluggish Market Share

6.1: Introduction

This chapter develops a model to show how the sluggish market share resulting from switching costs, together with *ex ante* imperfect information about market prospects, can lead to the exit of recent entrants. Previous models have demonstrated that locked-in customers constitute a credible entry deterrent and that real switching costs may support monopoly pricing post-entry. We add to these analyses by demonstrating that entrants may exit if they enter markets in search of new business but are unable to attract a sufficient share of the lucrative established customers, and find that the business they do attract is insufficient to cover their fixed or opportunity costs. We also find that our results are sensitive to the number of incumbents assumed. A larger number of incumbents implies more aggressive pricing on their part and the exit story is convincing in this case, since the entrant is then dependent on the new customers of lower value. Other models of exit have only considered monopoly incumbency.

Our consideration of entry and exit in the finance industry is dictated by the need to reconcile empirical observations with two theoretical expectations. Firstly, the tacitly collusive prices modelled in Chapter 4 would be expected to attract entry which should erode the market power exercised by price setters. Secondly, the literature on financial intermediation in LDCs maintains the hypothesis that inefficient intermediaries on these markets are protected by regulatory barriers to entry.

Shaw emphasized the disciplinary role of charter policy in enforcing competitive behaviour among financial intermediaries. This is in accord with the textbook model of a competitive industry: in the absence of scale economies or regulatory barriers, monopoly prices attract entry until prices are bid down to competitive levels. But, Shaw argued ([1973], p.88), existing "charter policy does not make the most of competition among banks and other financial institutions". It is the authorities' reluctance to issue operating licences that inhibits entry and permits incumbents to exercise oligopoly power in the repressed economy. However, models of strategic firm behaviour indicate that regulation is not the only entry deterrent: entry may also be deterred by

certain structural factors, one of which is switching costs. Observed behaviour in the Caribbean suggests consideration of structural entry barriers is insufficient. Entry appears to have been all too easy, with a record of entry, subsequent losses and failure and/or exit. The aim here is to provide an explanation of these observations.

Banks enter a market where they are at a disadvantage relative to incumbents in two respects. Current demand depends not only on current pricing but also on the market share resulting from past competition. In addition, new banks have fixed or opportunity costs of operation. Bank entrants acquire mainly new customers of uncertain value and switching costs imply that it is difficult to acquire market share among the established banking public. If the business provided by the new clientele proves insufficient, the entrant must exit. Our analysis has features in common with several of the many models in the literature which consider entry and exit behaviour. These are discussed in Section 6.2 below. Section 6.3 describes the Caribbean environment being modelled and the simpler, duopoly version of the model is described in Section 6.4. Appendix 6.1 presents the 3-firm version which is a better, but algebraically tedious, description of reality. Section 6.5 concludes.

6.2: Entry, Predation, Exit: A Survey

Introduction

While models of the potential for competition between firms have tended to concentrate on deterrents to the entry of potential rivals, there is increasing recognition that entry deterrence, predatory pricing and exit are closely connected. The behaviour analyzed under these categories reflect ways in which the relative strengths of long-established firms and their potential competitors interact, given the requirement of credibility for their possible strategies in response to rivalry. Relative strengths are expressed in terms of committed capital, reputation, relative costs (including advantages gained from experience) and the size of the expected profits to be defended in the future markets. The surveys by Gilbert [1988], Roberts [1987] and Wilson [1989] all treat these connections. A variety of models have considered conditions under which a) effective entry

deterrence is possible, b) entry may result in predatory pricing or wars of attrition, c) entry may be unsuccessful, culminating in exit. Switching costs have been analyzed as one of the fundamental features from which the behaviour described under a) and b) may result, but have not previously been considered as a cause of exit.

Entry Deterrence

Consider the key feature of entry deterrence: it depends on the existence of conditions that credibly impede capital mobility (see Gilbert's [1988] survey). The idea is that a market contains incumbents whose position is a given resulting from history or chance and who are able to impede entry by virtue of a first mover advantage. Credibility is crucial (and is captured by looking for a sub-game perfect equilibrium[1] - Selten [1975]) because rational potential entrants are not interested in the pre-entry situation but in the present value of post-entry profits, since only the incumbents' actions post-entry can affect entrants' profits.[2] Essentially, under complete information, strategic entry deterrence requires that there be some link between the incumbent's position pre-entry and the expected outcome after entry (Gilbert, op.cit.). It is this link that provides credibility. Variables that provide credibility include sunk costs which constitute an exit barrier for the incumbent, firm-specific scarce resources owned by the incumbent, cost-reducing experience and contracts which an incumbent may make with potential customers. Quality differentiation (which requires an investment whose benefits to an entrant are reduced by competition, so that customers are reluctant to experiment with his products) may also serve the same purpose, as may a market where it is costly for customers to switch suppliers.

The contracts and customer-costs-of-substitution cases have been dealt with in the switching costs models most extensively analyzed by Klemperer. Klemperer [1987b] demonstrates that very high or very low switching costs may deter entry because they act to lend credibility to

---

[1] Players choose their optimal strategy given the strategies of others (a Nash equilibrium) in every subgame so that in later periods no player has an incentive to change his strategy, it being optimal for the remainder of the game, i.e. for the subgame.

[2] This assumes complete information. Under incomplete information, pre-entry observations may be important to the potential entrant because they can provide information about the most likely post-entry outcomes.

post-entry strategies which reduce expected profits to the potential entrant. Increasing output (lowering price) pre-entry builds a high base of customers, deterring entrants who will have to pay their switching costs. On the other hand, low pre-entry output can signal that the incumbent will fight in the event of entry, this tendency is likely to be more important in a market where demand is growing or new customers are entering. As noted by Klemperer, op. cit., and Farrell [1986], as long as price discrimination is not possible, switching costs impose a certain ambiguity on competitive behaviour. The market share in locked-in customers tempts an incumbent to exploit these rents by pricing high (the 'fat cat' strategy stance[3]), but the noncooperative instinct induces him to fight by lowering price (the 'lean and hungry look').

Farrell and Shapiro [1988] analyze the fat cat effect of switching costs in a large market where there is repeated competition for old and new customers, with the incumbent acting as a Stackleberg leader in a pricing game. With no, or only moderate, scale economies, it is advantageous for the incumbent to exploit its old customers with high prices, rather than competing aggressively, so that the entrant is able to capture the new customers. As discussed in Chapter 4, switching costs act to enforce tacit collusion, but here the collusion takes the form of specialization in old customers (the incumbent) and new customers (the entrant).

The effect of switching costs on entry when the lock-in is governed by an explicit contract between buyer and incumbent is analyzed by Aghion and Bolton [1987] who find that the contract provides the incumbent and a single buyer with monopoly power over the entrant. It may both deter entry (since the entrant has to pay the customer's contract liquidation damages) and impose a social cost if the entrant is more efficient than the incumbent.

Predatory Pricing

In the entry context rational predatory pricing emerges as a means of lending credibility to threats of future action in a situation where the potential entrant has incomplete information about the market and/or the strength of the incumbent. A strategy is usually considered predatory if it

---

[3] This is Fudenberg and Tirole's [1984] terminology. See also Bulow, Geanokoplos and Klemperer [1985].

involves current losses designed to maximize future profits by discouraging future entry or inducing exit.

Klemperer [1986b] demonstrates that switching costs can produce a pattern of pricing that looks predatory. Immediate post entry price declines as the entrant captures, and the incumbent defends, market share. Price later rises as both entrant and incumbent exploit their locked-in customers. Similarly (Fudenberg and Tirole [1983]), where firms learn from experience, there is a strategic incentive to increase output (lower price) in the early phase of competition because, not only do later production costs decline as a result, but rivals' knowledge of the firm's lower costs may lead the rival to reduce later production, facilitating later price increases.

A lack of common knowledge is required for 'real' price wars. Where one firm has information on how profitable the market is likely to be for its rival the attempt to credibly signal disadvantageous conditions, and hence influence a rival's future behaviour, can result in lower prices as firms battle for market share (see Roberts [1987]). In the long run, however, firms do not deviate from the action they would have taken under full information because in equilibrium they are not fooled. The lowering of prices serves to permit credible separation between different types of signalling (informed) firms.

These results may all be interpreted as battles for market share. There is also a class of models where predatory action is able to alter rivals' behaviour. Thus Milgrom and Roberts [1982] show that, when an entrant has incomplete information about the incumbent's preferences, costly predation may reinforce a potential entrant's beliefs that the incumbent prefers predation. It may thus deter potential entrants with opportunities for returns elsewhere that exceed the return with predation. Kreps and Wilson [1982] also derive a similar result. The incumbent is here investing in reputation whose present value exceeds the short-run costs of predation.

Exit

Relative to entry models there are not many models of exit. It is not uncommon for these to aim at explaining an entrant's exit since, although empirical observation has suggested that large incumbents may force entrants out (see Scherer [1980], Chapter 12), the theoretical expectation is

that the entrant, foreseeing an incumbent's losses from predatory pricing, will be prepared to wait him out. None of these specifications capture our empirical story.

Entry-then-exit would be expected in contestable markets (Baumol [1982]) where free mobility and the absence of sunk costs enforce marginal cost pricing. Potential rivals enter to take advantage of a profitable opportunity and then exit. The entry-exit phenomenon observed in the Caribbean is not consistent with contestability since entry there has often resulted in losses, whereas the contestable markets' entrant is motivated by profit[4].

In Benoit's [1984] model with complete information, the entrant may leave if financially constrained to withstand only a finite period of predatory pricing. The threat to fight induces immediate exit, and with positive entry cost there is no entry, since the fighting incumbent will eventually induce exit. This model produces a better explanation of the deterrence of small entrants than of entry with exit.

Asymmetries of information are usually required for a convincing picture of entry with exit. Benoit, op.cit., models an incumbent who is unsure of the financially constrained entrant's commitment to staying in. Entry sometimes occurs, followed by predatory pricing which may end in cooperation or exit depending *inter alia* on the probability the uncommitted entrant assigns to staying in. The existence of financial constraints may explain cases of entry and failure of small intermediaries in the Caribbean but the asymmetry in financial constraints is not attractive for large multinationals.

Roberts' [1986] model also invokes incomplete information to generate predatory pricing and exit after entry. In a situation where the incumbent has private information about the strength of market demand and his quantity choice is unobservable, the incumbent has an incentive to signal low demand by increasing output. The entrant infers demand from the resulting market price. A credible signal requires an otherwise-unprofitably-high quantity level (low price). In a separating equilibrium (where the information can be inferred from the signal - the only one considered

---

[4] Of course, our postulates of switching costs and imperfect information are also inconsistent with contestable markets, but as hypothesized priors they cannot serve to refute the possibility that the Caribbean banking market is contestable. The empirical observation of losses does so.

by Roberts) the entrant makes the correct inference and exits only when demand is in fact low. Thus predation does occur but does not induce more exit than would have taken place under complete information. It is the entrant's learning that demand is low (as in our model), rather than asymmetric information, that explains exit here. However, entry with exit depends on the non-observability of quantity choice, which does not appear a good representation of the financial industry where price rather than output is the strategy variable.

These models consider entry onto an 'occupied' market where the entrant is at a disadvantage relative to the incumbent, a structure similar to the empirical structure we model. Fudenberg and Tirole [1986] consider a different structure where two firms enter simultaneously, each with private information on its own opportunity or fixed costs, and exit time is the strategy variable. There is dynamic selection game (war of attrition) in which the firms stay in until one becomes more pessimistic about the strength of its rival, with the higher opportunity cost firm exiting first. If duopoly profits increase fast enough, both firms may remain in. In contrast to the growing market, Ghemawat and Nalebuff [1985] model exit in a declining industry. They find that the largest firm leaves first, unless there are strong economies of scale. It is evident that neither of these models would be a good description of our stylized facts (see Section 6.3) but, like Fudenberg and Tirole, we assume that fixed or opportunity costs are the basic cause of exit.

In Section 6.4 we show that entry with exit can be explained with the simpler assumptions of sluggish market share and symmetric uncertainty about the volume of new demand if incumbent market share is not so great as to limit aggressive pricing and the entrant has fixed or opportunity costs.

6.3: Entry and Exit in the Financial Markets of the Caribbean

Entry and exit in the four Caribbean financial markets being considered have essentially been free of regulatory constraints. Most of the large commercial banks are multinationals which predate the current central banks, and their global size relative to the Caribbean economies is such as to make their domestic failure a remote possibility. As a result, the regulators have not

had, until recently, to be concerned with the protection of depositor motive for restricting entry. Banking licences have therefore been issued to any international bank of good standing. Further, far from wishing to protect the interests of established local interests - the usual reason assumed for entry-inhibiting charter policy - the regulators have viewed new international banks as providing additional access to much-needed foreign exchange resources. Of the four countries, the only one to impose any 'rules' which could deter foreign banks was Trinidad and Tobago which, as a petroleum exporter, perhaps considered itself rather more immune to balance of payments constraints. However, these rules do not affect domestically owned institutions. In Chapter 4 we argued that domestically owned competitors, in the institutional form of finance houses, did not provide effective competition for banks in deposit-taking because of the riskiness attached to their deposits. The perception of that risk is based not only on their prior lack of reputation but on potential depositors' observations of their exit from market, following failure - the reputation acquired proved to be negative. In this context, then, we should also consider such finance institutions as have tried to establish themselves as potential competitors with commercial banks.

The Bahamas' (see Ramsaran [1984], Ch.2) regulations, fixed in reaction to fraudulent operations prior to the central bank's establishment, require a banking licence and the publication of operating statements, and establish the right of supervision. No capital asset requirements are specified. In Barbados, it is only as a result of the recent establishment of regionally-owned banks or locally-owned finance houses, that the authorities have begun the process of revising banking legislation in order to lay down capital requirements for banks. Previously, capital requirements were not fixed for the branch banks[5] which dominate the banking sector. With an obviously prudential motive, existing legislation mainly constrains advances to insiders, restricting loan size to single customers and equity participation in domestic non-bank companies to 25% of paid-up capital and reserves, as well as providing the authority for inspection and monetary policy measures (see Central Bank of Barbados [1986]). Similarly, entry regulations in Belize do not impose capital requirements. Legislation provides for prudential rules restricting loans

---

[5] Branch banks have no legal status separate from the parent company.

to any single customer from exceeding 25% of bank capital (Luben [1983]). None of these regulations seem likely to deter entry - even the prudential measures may have little force because the capital and paid up reserves refer to those of head office liabilities in the case of multinational banks. In effect, the regulatory authorities have relied to a considerable extent on the banks' management capabilities and the discipline of reputation maintenance in dealing with the multinationals as well, of course, as on the supervision of the banks' national authorities. The situation in Trinidad and Tobago differs somewhat because most multinational banks assumed subsidiary status[6] after 1970 when government required that existing banks have a minimum of 51% local equity participation. When "localization" took place in 1972, capitalization share issues were made to meet capital requirements.

In general, there are capital and reserve fund requirements for locally or regionally incorporated banks but these are not fixed to deter entry. There are also non-regulatory costs of establishment. The major item of set-up costs would be an initial loan from headquarters. While this may represent only "seed" money, in view of the size of the multinational banks' assets relative to the economies of the islands, it remains an opportunity cost of operating in a particular economy until adequate local deposits have been acquired. The only sunk costs may be pension funds etc. for employees. It should perhaps be mentioned that exchange control regulations permit the automatic repatriation of both after-tax profits and "shares in head office administrative expenses".

It would appear that the entry decisions of the commercial banks are made on regional and/or global strategic grounds, the domestic economies of the individual countries having little influence. For example, following the Commonwealth Sugar Agreement and International Sugar Agreement of 1953, prescribing sugar export quotas, the Bank of Nova Scotia entered the region to take advantage of the resulting boost to economic activity. Citibank entered Trinidad and Tobago in 1965 and Barbados in 1968, during its worldwide expansion. More recently, the Bank of Credit and Commerce International, apparently in the process of expanding into the Latin

---

[6] That is, they acquired a legal identity separate from the head office bank.

American market, opened offices in Miami, Jamaica and Barbados, and is in the process of nego-

tiating with the Trinidad and Tobago government.

The ease of entry and exit is demonstrated by the experience of Barbados. Towards the end

of the sixties there were four long-established banks (Barclays, Royal Bank of Canada, Canadian

Imperial Bank of Commerce and the Bank of Nova Scotia). On the US banks' global expansion

wave of the sixties and seventies, Citibank (1968), the Bank of America (1969), Chase Manhattan

(1971) and the Bank of Chicago (1974) entered. However, none of these newer entries were very

long-lived. It has been hypothesized that the major factor in their exit decision was their failure

to attract sufficient large-scale depositors, and the concomitant business, to make their operations

viable. The only one of them with some success in this area, Chase Manhattan, was the longest-

lived, exiting only in 1986. Both the Bank of America and the Bank of Chicago exited in 1978.

Their portfolios were taken over by the government-created Barbados National Bank (BNB),

benefiting from the substantial deposits of the Barbados Savings Bank, which formed the basis of

the BNB's Commercial Division. Citibank left in 1984, its assets being purchased by the

newly-formed Caribbean Commercial Bank (owned by a Trinidad and Tobago-incorporated

insurance company) which later also bought Chase Manhattan's assets. In 1983 there was a new

multinational entrant: the Bank of Credit and Commerce International. Whether or not these

latter remain is also thought to depend on whether they will be able to attract sufficient custo-

mers.

Except for a relatively brief period between 1973 and 1974, the entry of these banks does

not appear to have had a substantial effect on the pricing behaviour of the banking sector. In

1973 there was some attempt to actively compete on posted deposit rates but these were frus-

trated by official action. It could be argued that the official regulation of interest rates in Barba-

dos throughout this period had a determining influence on the banks' exit decisions and

behaviour. Two facts are inconsistent with this argument. Firstly, legal ceilings on interest rates

had been in effect in Barbados for many years (for example, the ceiling was raised from 6% to

8% in 1961) and forward-looking banks would have taken this into account in planning post-

entry strategies. For example, the Bank of Chicago entered after the reimposition of ceilings.

Secondly, the ceilings were usually several percentage points above actual rates. The ceilings became effectively binding only in October 1973 when they were removed, to be reimposed, although at higher levels, when the new central bank saw the resulting jump in rates. Nor does the flurry of rate competition experienced in 1973 (see Howard [1976]) appear to have been greatly influenced by domestic battles for market share. The rise in rates at that time mirrored the global rise in bank rates resulting from world-wide inflation. The average three-month euro-dollar rate rose from 5.94% at the end of 1972 to 10.99% in the third quarter of 1973 at a time when the banks had heavy commitments on the eurodollar markets. By 1976, the ceilings again ceased to be effective constraints on rate movement.

The entry of non-bank competitors is illustrated by the Trinidad and Tobago experience. The petroleum boom in the 1970s increased disposable income and liquidity in the Trinidad financial market but commercial bank loan rates did not fall (Bourne [1985]). High liquidity and increased demand for housing construction and consumer durables provided an incentive for the entry of finance houses. Between 1970 and 1978 finance houses' share in the financial assets of the household sector grew from 0.6% to 2.3%. These institutions were also subject to looser regulatory constraints, with a legal reserve requirement of only 3% (banks' requirement was 14%) and closer ownership connections with the corporate sector than banks are permitted. However, far from serving to stimulate competition, their relatively unimpeded entry into the market has resulted only in failure of several of these companies. The only available analysis of these failures is a treatment which falls within the ambit of political science rather than economics: Parris [1984] assesses the events surrounding the failure of one of the largest of these finance houses, rather than the conditions which produced failure. However, Parris characterises the process as a 'game being fought for control of the financial system in the country' between the 'old financial elite' and the 'post-colonial political elite'. In less rousing terms, his discussion suggests that the regulators' wish to encourage a more competitive financial system was frustrated by the dominant banking firms' oligopolistic control of financial resources. The run on deposits, which started when the finance house and its holding company reported a loss, could have been stemmed

'... with financial support from the old financial elite, (this option), though acceptable and possibly the option most preferred by the Central Bank, was obviously unacceptable to the old financial elite, whose goal was to eliminate these "fringe" institutions.' (Parris, op.cit., p.25)

A 'shallow pocket' type of explanation may apply: newer institutions, who are more vulnerable because of their smaller, less valuable customer base may require substantial financial resources in order to withstand adverse factors. Then dominant banks, who provide considerable support to each other in terms of interbank loans, can restrict rivalry to the oligopolistic group simply by refusing to extend similar facilities. However, Parris's discussion does not rule out the possibility that the institution was ineptly (unethically?) managed and hence not a viable proposition for rescue, unless administratively restructured.

Other failures of similar institutions have occurred in Trinidad and Barbados. The explanation proposed here is based on the idea that these firms see large profits of existing banks, (or, in the case of Barbados, the restrictions on banks' credit operations) as providing an opportunity for profitable operation. This may in part be fostered by the looser regulations. However, the established banks are able to hold onto their market share, and the new entrants' inability to capture substantial volume makes them highly dependent on the value of the business obtained from initially unattached customers, which they cannot observe until they start operating. If this is not sufficiently lucrative, it does not cover their fixed costs, the resulting losses scare those customers obtained, and failure follows.

6.4: A Model of Bank Entry and Exit

Background

Our empirical context has suggested that there are market features relevant to an explanation of entry-with-exit which are not considered in existing exit models. For example, (a) switching costs have not previously been considered as an influence on exit, (b) models usually require an asymmetry of information to give exit, (c) only duopolistic post-entry competition is considered. With reference to (c) a point of interest is that the plausibility of exit under our assumptions depends on the numbers already in the market, which determines the degree of aggression displayed by incumbents. This aspect of potential competition is not captured in models which consider only monopoly incumbency - all other models in fact.

We wish to explain why potential rivals to long-established intermediaries in the banking market of the Caribbean have been unsuccessful entrants, testing the idea that their failure to establish themselves resulted from their inability to capture a sufficient market share of the most valuable customers on the market. The 'test' consists of discovering whether we can show that, with switching costs producing sluggish market share, a noncooperative equilibrium exists in which banks enter the market, only to exit if they learn that the business provided by the clients most readily acquired is of low volume. The assumptions of an uncertain volume of new business whose realization is learnt at a cost are consistent with the incomplete information environment McKinnon Shaw described for LDCs.

Assumptions

1) We model market growth by the entry of new customers whose business volume is less than or equal to that of old customers. Switching costs require that old and new customers be distinguished. Therefore, Farrell and Shapiro [1988] consider new customers, but view new and old as homogeneous, except in attachment. But, unlike a one-shot business, banks are concerned with customer prospects - the distinction between new and old customers arises in part from the uncertain value of new customers, a distinction made by credit rating agencies, for example.

2) It is assumed that banks are faced with a net demand for services and set a single net price. Although it would be desirable to model banks as rate setters in both their input and output markets, a model incorporating this feature, as well as entry behaviour and uncertainty, is unnecessarily complicated. In addition, deposit and loan rates are not set independently. Thus, while recognising that important aspects of strategic behaviour may be obscured, we may view banks as setting their rates in tandem, the loan rate being determined as a markup over the deposit rate to cover the marginal costs and risks of credit. It is equivalent to the usual assumption of price-setting with zero marginal cost.

3) The demand function assumed is a linear version of that in Farrell [1986], who used the general form to indicate how the inertia created by switching costs influences the one-period problem of two established firms. Current demand of firm i in period t with two firms, i and j is given by:

$$h_{it} = h_{it-1} + p_{jt} - p_{it}$$

where h_{it-1} is the market share inherited from the last period, and current demand only exceeds original share if current price is less than the rival's. The price differential captures sluggish price response. This function has the linear form usual in models of differentiated product markets (see Friedman [1986]), recognizing that switching costs serve to differentiate an otherwise homogeneous market. It illustrates the conflicting incentives of firms in markets with switching costs when price discrimination is ruled out: to price low to capture additional customers and to price high in order to exploit their existing market share. Those with the highest market share have the most to lose by lowering price and hence will tend to price higher. Total industry demand is assumed inelastic.

4) It is assumed that new customers allocate themselves to banks in inverse proportion to existing market share. In this way we allow for customer recognition that high customer base banks will set the highest prices.

5) We attempt to capture a pricing feature peculiar to banks. Most product market models assume single product firms so that each competitor need fix only one price. However, single-price determination, on even one side of its market, is obviously at variance with bank practice. Banks usu-

ally set a range of linked interest rates. Deposit rates may be fixed by term and/or size of deposit, reflecting the fact that customers fall into segments which may be determined by income and opportunity, among many other possibilities. The introduction of new customers basically segments the market served by the banks. As a result, banks fix prices which, given the additive separability of the demand functions, are themselves additively separable in customer types. We can therefore regard a bank as determining a price structure whose components depend on the type of customer the bank has, or wishes to attract.

6) We assume that the entrant has a periodic fixed cost of operation, or an opportunity cost (compare Fudenberg and Tirole's [1986] model). Given the net price strategy used, the entrant must have a cost of operation which is independent of scale in order for exit to be a possibility. But the assumption, and the asymmetry it induces between entrant and incumbent, is quite realistic when considered in the light of our stylized facts. The relative positions of long-established and newly-arrived financial institutions do differ, especially when both institutions are multinationals owned outside of the domestic economy. The localised incumbent operates with already trained staff and domestic management and owns their infrastructural facilities.[7] Their familiarity with the economy and market participants implies that they already have the base (that is, non-specific to particular loans) economic information necessary to undertake loan allocation. On the other hand, the entrant must maintain expatriate management, paying assignment costs and staff training costs that do not vary with the volume of business undertaken. It also has to acquire the basic, and costly, information required for financial intermediation in the specific economy. In addition, if, as argued in 6.3, multinational banks tend to enter particular regions on global strategic grounds, they would need to consider whether the opportunities foregone by operating in a particular economy are justified by the profits there. By virtue of their enduring incumbency, one could presume that long-established intermediaries have already decided on the particular economy.

---

[7] Thus the Central Bank of Barbados [1986] green paper proposes that such facilities could constitute the capital fund for already-established banks.

The market game is viewed as having two stages. A new bank and new customers enter a market with one (or two in the appendix) established bank(s) at the beginning of the first (entry) stage, the new bank attracting an above average share of new customers. However, the magnitude of the business provided by the new customers is not known to the banks. Incumbent(s) and entrant fix prices simultaneously, taking account of the effect current price has on future market share and hence profits. It is only after observing profits at the end of this stage that the banks are able to assess the value of the new customers. If this value is high enough to cover its opportunity costs, the entrant stays in. If not, it exits and the incumbent(s) acquire its customers and is a monopoly (compete as a duopoly) in the second stage. The one incumbent/one entrant version of equilibrium behaviour is described below. In this the entrant acquires all the new customers so that the incumbent must infer their value from the entrant's exit decision (he has not had the opportunity to assess their value himself).

The model was initially formulated with two incumbents and one entrant because, given our priors about the importance of established market share, the concentration of share resulting from monopoly incumbency seemed likely to misrepresent empirical strategic behaviour. The algebra of triopoly is cumbersome in this linear model, however, and is not necessary to an intuitive understanding of the model structure. The triopoly model has therefore been largely relegated to Appendix A6. However, in line with our initial intuition, the results of the triopoly version do provide a more plausible picture for exit. The interfirm elasticity of demand means rivalry among fewer firms, especially between two firms only, is diluted - the entrant acquires a higher share of established customers because of higher incumbent prices. The fewer the incumbents, the fatter the cat[8] and the less incentive there is to lower price. (See also assumption 3) above). Not only does the monopolist's size encourage price exploitation but a high share acquired by the entrant increases the latter's second period price and reduces second period rivalry if the entrant does stay in. This suggests a type of collusive tendency.

Considered from the viewpoint of the global market place, the model is rather akin to a

---

[8] To use Fudenberg and Tirole's, op.cit., expressive terminology.

search model, although we are looking at decisions within one economy (store) only (in the short-run). Expected losses for one period are the costs of acquiring information, given a strategy of locating the most profitable markets. In the long-run within a region (presuming that there is a known distribution of customer values by region) banks' expectations will prove correct and they will locate an economy appropriate to their fixed costs.

Model

Consider a market where customers acquire switching costs once they have opened accounts and banks with higher market shares charge higher prices. A bank entrant into a monopolistic market will acquire all the new customers (see assumption 4))[9]. Therefore his expected demand in the entry period is

$$h_{E1} = N^e + p_{i1} - p_{E1} \tag{6.1}$$

whereas the incumbent has demand

$$h_{i1} = L + p_{E1} - p_{i1} \tag{6.2}$$

where h_{E1} is the demand faced by E in period 1, N_e is the expected volume of new business, p_{i1} is the incumbent's price in period 1. L is the volume of established business which we normalize to have value unity.

Nature chooses N according to distribution ρ which gives the banks' prior beliefs. Banks are not informed of the chosen N. It is assumed that incumbents and entrants have common beliefs. Banks cannot assess the 'value' of new customers until they have had the opportunity to observe the net revenues resulting from operation of their accounts after first period pricing. The expected volume of business provided by new customers is

$$N^e = \rho N^a + (1 - \rho) N^b$$

where $N^b < N^a \leq L = 1$ and ρ is the prior probability attached to the occurrence of N^a. We assume that new customers enter the market only in the first stage of the two-stage game.

---

[9] In appendix equation (A6.1) $K = 2$, $l_{E0} = 0$ and $l_{i0} = 1$.

E, the entrant, has a fixed or opportunity cost, F, of operation, in each period and we assume that this is such that E prefers to leave the market if it learns that $N = N^b$. That is, we are assuming that

$$\Pi_{E2}^{d}(N^b) \leq F < \Pi_{E2}^{d}(N^a) \tag{6.3}$$

where the superscript d refers to duopoly equilibrium values[10]. $\Pi_{E2}^{d}(N_b)$ is thus E's profits in a second stage duopoly when $N = N^b$. This exit rule is known to the incumbent banks.

Equilibrium

The equilibrium concept used is sequential equilibrium (Kreps and Wilson [1982]). This requires that a player's system of beliefs and strategy (its assessment) be both consistent and sequentially rational. Sequential rationality requires that the strategy of each player starting from each information set be optimal starting from there according to some assessment over the nodes in the information set[11] and the strategies of the other players (Kreps and Wilson [op.cit.], p.87). The consistency condition is that the assessments obey Bayes' rule so that a player's beliefs must be consistent with its knowledge of the structure of the game and its knowledge of other players' strategies, satisfying rational expectations.

A strategy for firm k specifies: i) a first stage price, p_{k1} , ii) for the entrant an exit rule according to the realized value of N at the end of stage 1 and iii) stage 2 prices from the realization of stage 1 demand and the exit decision. An equilibrium occurs when each bank maximizes expected discounted profits given that its rivals are following their hypothesized strategies. Where information about N is symmetric but incomplete, prices are chosen with no direct information about the magnitude of N or other banks' choices. We assume that the banks do not directly observe N but must infer this from profits. That is, banks require a period of 'learning' to assess new clients. With a monopoly incumbent who has no opportunity to learn about the new customers directly, their value is inferred from E's exit decision. At the beginning of the second

---

[10] Note that small d is used to indicate duopoly in the model with monopoly as an alternative. In the model of the Appendix, where the two possible market structures are duopoly and triopoly, capital D is used for duopoly values.

[11] The information set is the set of decision nodes among which the player cannot distinguish, i.e. the player does not know at which precise decision point he is in fact located among the possible set.

period, banks have observed the previous period's demand and prices. In period 2, therefore, banks know the value of N. Each bank chooses price to maximize period 2 profits conditional on its period 1 observation of demand and the exit decision. In period 1 each chooses the price that maximizes expected discounted profits given its period 2 decision rule and its hypotheses concerning the behaviour of rivals, particularly the entrant's exit rule.

Equilibrium behaviour

Only equilibrium behaviour for a monopoly incumbency is analyzed here, although comparisons are made with the results of the duopoly incumbency model given in Appendix A6. As usual we begin by considering stage 2. In equilibrium E will know the magnitude of N from his first period profits. If E exits his customers move to incumbent. If E remains in the market the banks compete as a duopoly.

If $N = N^a$, the entrant remains in the market and chooses p_{E2} to maximize:

$$\Pi_{E2} = p_{E2} h_{E2} = p_{E2} [h_{E1}(N^a) + \bar{p}_{i2} - p_{E2}] \tag{6.4}$$

where $\bar{p}_{i2}$ is the price E conjectures that i will set. $h_{E1}(N^a)$ is the demand resulting from pricing in Stage 1 when $N = N^a$. Since $N = N^a$ and E has remained in, i maximizes:

$$\Pi_{i2} = p_{i2} h_{i2} = p_{i2} [h_{i1}(N^a) + \bar{p}_{E2} - p_{i2}] \tag{6.5}$$

The first order conditions give best response functions:

$$p_{E2} = \frac{1}{2} [h_{E1}(N^a) + \bar{p}_{i2}] \quad and \tag{6.6}$$

$$p_{i2} = \frac{1}{2} [h_{i1}(N^a) + \bar{p}_{e2}] \tag{6.7}$$

These best response functions are illustrated in Figure 6.1

Figure 6.1:

Reaction Functions for Stage 2 Duopolistic Competition
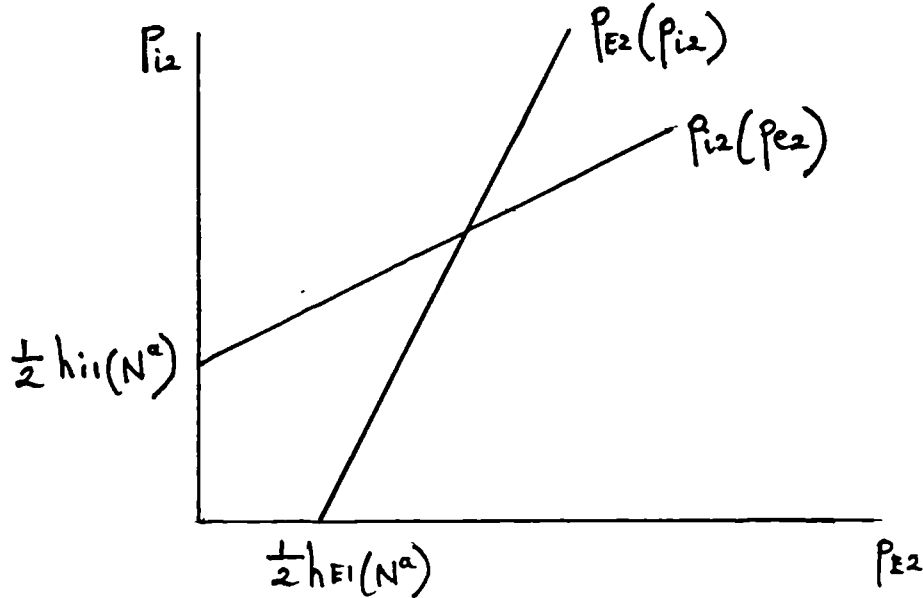
In a Nash equilibrium where each firm take the other's prices as given, simultaneous solution of the best response functions give equilibrium duopoly prices:

$$p_{E2}^d = \frac{1}{3}[2h_{E1}(N^a) + h_{i1}(N^a)]$$ (6.8)

where the equilibrium price is symmetric for incumbent i. Equilibrium price depends positively on the market shares of all banks, with greater weight given to own share. $p_{E2}^d > p_{i2}^d$ if $h_{E1}(N^a) > h_{i1}(N^a)$, that is, if E's inherited market share is larger, and conversely. In the final period since higher market share provides an increased incentive to price high, bank services are strategic complements [12] and banks' prices are increasing in all shares. Substituting prices of the form in (6.8) for i and E into the demand functions in (6.4) and (6.5), evaluated at equilibrium prices, realized demand for E in stage 2 is

$$h_{E2}^d = \frac{1}{3}[2h_{E1}(N^a) + h_{i1}(N^a)]$$ (6.9)

---

[12] B regards A's product as a strategic complement if B responds to an increase in p^A by raising p^B. If p^B is decreased, the products are strategic substitutes: Bulow, Geanakoplos and Klemperer [1985].

It is evident that if i begins the second period with a higher market share and hence prices higher in period 2, he loses customers to E. E's second period profits are, from (6.8) and (6.9):

$$\Pi_{E2}^d = \tfrac{1}{9} [2h_{E1}(N^a) + h_{i1}(N^a)]^2 \tag{6.10}$$

These are the second period profits which E will expect, and symmetrically for i, when he has learnt $N = N^a$ at the end of period 1. In order to find the effect on second period profits of first period market share and hence first period pricing, substitute for h_{E1} and h_{i1}, evaluated at $N = N^a$ from (6.1) and (6.2) into the profit functions of (6.10). Duopoly second stage profits for E in terms of stage 1 prices are

$$\Pi_{E2}^d = \tfrac{1}{9} [2N^a + L + p_{i1} - p_{E1}]^2 \tag{6.11}$$

and for i they are

$$\Pi_{i2}^d = \tfrac{1}{9} [2L + N^a + p_{E1} - p_{i1}]^2 \tag{6.12}$$

If E has exited at the end of stage 1 because $N = N^b$, i will inherit all his customers and since market demand is inelastic he will set the reservation price and obtain monopoly profits which we denote by $\Pi_{i2}^M(N^b, L)$. E's profits in the particular market considered are zero.

Consider the first period. The incumbent knows that E has entered and its exit rule. The distribution of N is common knowledge. Current price influences current demand as well as the volume of business that the banks will have in the second period. The entrant therefore chooses first period prices to maximize expected discounted profits:

$$E\Pi_E + F(1+\delta\rho) = E\Pi_{E1} + \delta\rho\Pi_{E2}^d(N^a)$$

$$= p_{E1}[N^e + p_{i1} - p_{E1}] + \tfrac{\delta\rho}{9}[2N^a + L + p_{i1} - p_{E1}]^2 \tag{6.13}$$

where δ is the common discount factor, and ρ is the probability that $N = N^a$, that is, that E will remain in the market and compete with i in the second period. The first order conditions (which are necessary and sufficient for this problem) give the best reply function for E:

$$p_{E1} = [2(9-\delta\rho)]^{-1} [(9-4\delta)\rho N^a + 9(1-\rho)N^b - 2\delta\rho L + (9-2\delta\rho)\bar{p}_{i1}] \tag{6.14}$$

Similarly, i maximizes expected discounted profits over the two periods:

$$E\Pi_i = E\Pi_{i1} + \delta p\, \Pi_{i2}^d(N^a) + \delta(1-p)\Pi_{i2}^M(N^b)$$

$$= p_{i1}[L + p_{E1} - p_{i1}] + \frac{\delta p}{9}[2L + N^a + p_{E1} - p_{i1}]^2 + \delta(1-p)\Pi_{i2}^M(N^b) \qquad (6.15)$$

i's best reply function is therefore:

$$p_{i1} = [2(9-\delta p)]^{-1}[(9-4\delta p)L - 2\delta p\, N^a + (9-2\delta p)\bar{p}_{E1}] \qquad (6.16)$$

In contrast to the second period best replies (compare (6.6) and (6.7)), we note that a rival's inherited business volume has a negative effect on the optimal price of each bank. E starts the period with N customers and the first derivative of i's price with respect to N^a is negative, while i starts with L customers and the first derivative of E's price with respect to L is negative. Whereas consideration of only the single period problem illustrates only one aspect of the incentives in markets with switching costs - that of pricing high to exploit customer base - the forward-looking price reflects the incentive to price low in order to build a customer base for the next period. The higher the current share of rivals, the more there is to be gained from such pricing and hence price is negatively related to rivals' share. These conflicting influences on first period pricing are illustrated in Figure 6.2's three panels. Panel 6.2a illustrates the upward pressure on prices of the accommodating incentive given by own-share, omitting others' market segments from the intercepts of the reaction functions. Panel 6.2b illustrates the competitive tendency by ignoring the own-share effects on the intercepts and panel 6.2c shows the actual reaction functions which are the vertical (horizontal) sum of the functions in panels 6.2a and 6.2b.

Solving (6.14) and (6.16) simultaneously, and simplifying, we obtain the equilibrium first period prices:

$$p_{E1}^d = [6(9-\delta p)(27-4\delta p)]^{-1}\left\{p\left[3(9-4\delta)(27-4\delta p) + (9-2\delta p)[3(9-2\delta) - 4\delta(6-\delta p)]\right]N^a \qquad (6.17)\right.$$

$$+ 3(1-p)[9(27-4\delta p) + (9-2\delta p)^2]N^b$$

$$\left. + \left[3(9-2\delta p)[27-2\delta p(9-\delta p)] - 6\delta p(27-4\delta p)\right]L\right\}$$

Figure 6.2: *Reaction functions for Stage 1 duopolistic competition*

Panel 6.2c

$$P_{iI}$$

$$P_{iI}(P_{EI})$$

$$P_{EI}(P_{iI})$$

$$\frac{(q-4\delta\rho)L - 2\delta\rho N^{\alpha}}{2(q-\delta\rho)}$$

$$P_{EI}$$

$$[2(q-\delta\rho)]^{-1} [\rho (q-4\delta)N^{\alpha} + q(1-\rho)N^{b} - 2\delta\rho L]$$

Panel 6.2b

$$P_{iI}$$

$$P_{EI} (P_{iI}) \big|_{N=0}$$

$$P_{iI} (P_{EI}) \big|_{L=0}$$

$$-2\delta\rho [2(q-\delta\rho)]^{-1} L$$

$$P_{EI}$$

$$-2\delta\rho [2(q-\delta\rho)]^{-1} N^{\alpha}$$

Panel 6.2a

$$P_{iI}$$

$$P_{iI}(P_{EI}) \big|_{N=0}$$

$$P_{EI}(P_{iI}) \big|_{L=0}$$

$$\frac{q-4\delta\rho}{2(q-\delta\rho)}$$

$$\frac{\rho (q-4\delta)N^{\alpha} + q(1-\rho)N^{b}}{2(q-\delta\rho)}$$

$$P_{EI}$$
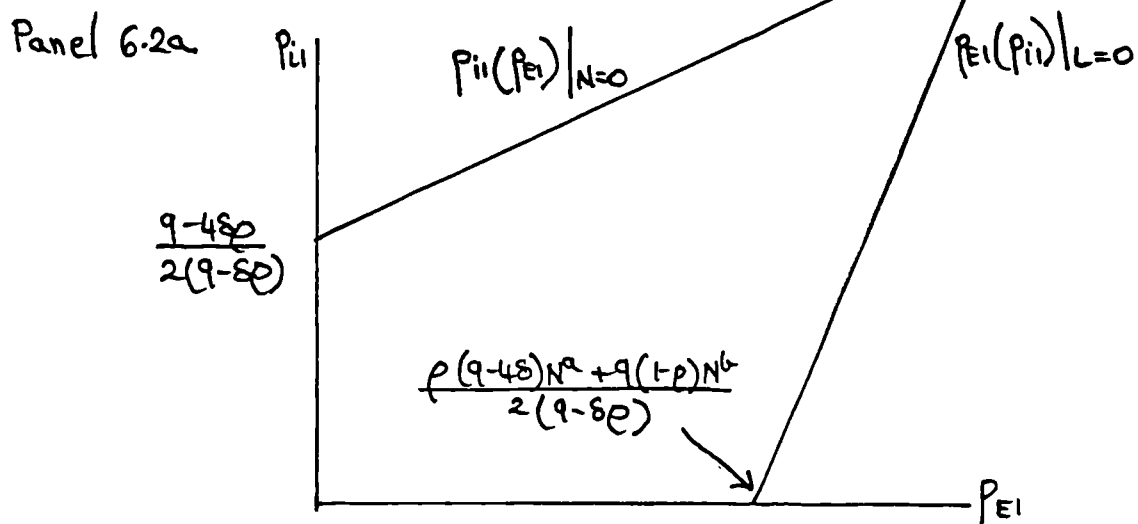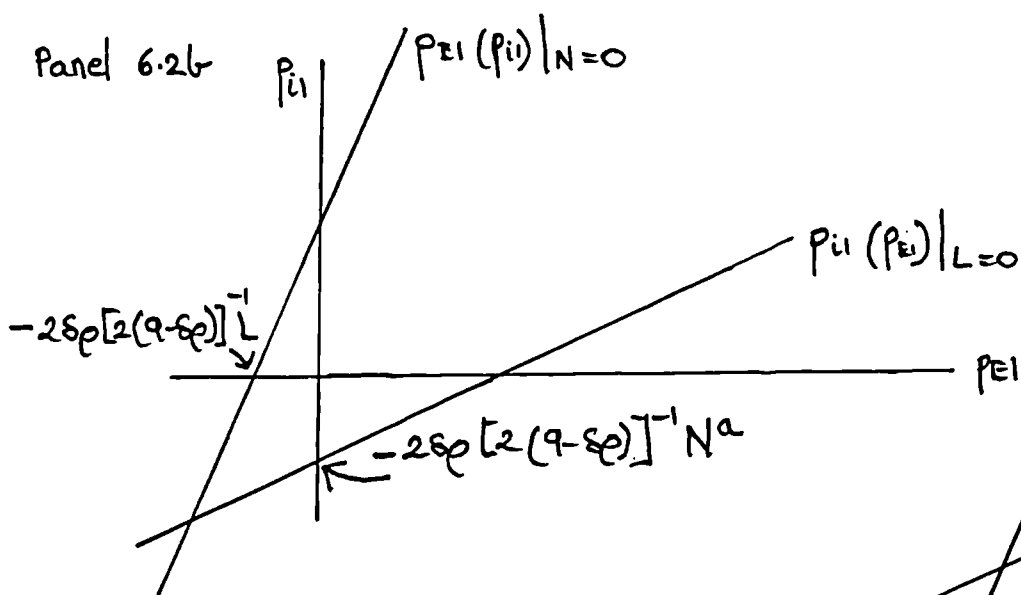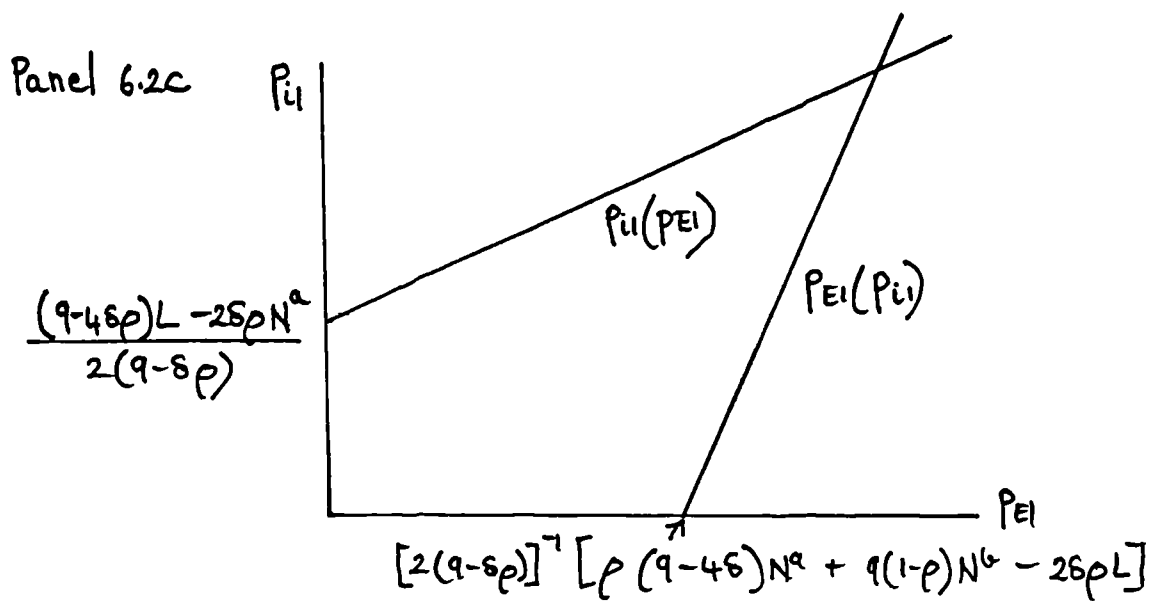
$$p_{i1}^d = [3(27-4\delta\rho)]^{-1}\left\{ 2[27-2\delta\rho(9-\delta\rho)]L + \rho[3(9-2\delta\rho)-4\delta(6-\delta\rho)]N^a + (1-\rho)(9-2\delta\rho)N^b\right\} \quad (6.18)$$

We may note that components of net price are always greater for the market segment in which a bank 'specializes'. The differences between p_{i1}^d and p_{E1}^d are given by

$$p_{i1}^d - p_{E1}^d = (27-4\delta\rho)^{-1}[(9-2\delta\rho)L - \rho(9-2\delta)N^a - 9(1-\rho)N^b] \quad (6.19)$$

For the L segment p_{i1} exceeds p_{E1} and vice versa for N. We would expect the entrant to price lower overall given its need to attract reliable clientele, and this is indeed so in the single period problem. In a single period game $(\delta = 0)$, $p_{i1}^d - p_{E1}^d = (1/3)(L - N^e) > 0$ However, in the two stage game E's incentive to offer lower prices is reduced by the possibility of exit before it could gain fully the returns on those prices. A large probability of N^b increases E's price (except for the unlikely N^a component).

Pricing is less competitive in the duopoly first period game. In the triopoly model other banks' market share have a negative influence on equilibrium price (compare these prices with those in equations (A6.19) and (A6.20)). Here they do not. The incumbent's monopoly of known, established customers makes exploitation of market share more attractive than in a duopoly. However, differentiation of (6.19) and comparisons with the effect of ρ on the two prices separately indicates that at higher ρ, when E is more likely to stay in, and expected N more valuable, i prices more competitively.

As a result of these prices, in period 1 banks obtain new business which varies for E according to whether realized N is N^a or N^b. Substituting p_{i1} and p_{E1} from (6.18) and (6.17) into (6.1), where N is evaluated at N^a for (6.20), and at N^b for (6.21):

$$h_{E1}^d(N^a) = (27-4\delta\rho)^{-1}[(27-9\rho-2\delta\rho)N^a - 9(1-\rho)N^b + (9-2\delta\rho)L] \quad (6.20)$$

$$h_{E1}^d(N^b) = (27-4\delta\rho)^{-1}[(18+9\rho-4\delta\rho)N^b - \rho(9-2\delta)N^a + (9-2\delta\rho)L] \quad (6.21)$$

Since i does not begin the period with N customers he has no opportunity to evaluate them in the first period and must infer their value from E's decision. He obtains first period business

$$h_{i1}^d = (27-4\delta\rho)^{-1}[2(9-\delta\rho)L + \rho(9-2\delta)N^a + 9(1-\rho)N^b] \quad (6.22)$$

Summing over these demands gives total market demand: $h_{E1}^d (N^a) + h_{i1} = L + N^a$ and $h_{E1}^d (N^b) + h_{i1} = L + N^b$, so that by specifying values for ρ and δ we may obtain relative shares. It turns out however that the value of δ makes little difference so we consider these shares with $\delta = 1$[13]. Although E's price for N exceeds i's, E keeps most of its new customers, whatever the value of N and ρ. In contrast to the case with entry period triopoly, E obtains almost a third of the L customers, as illustrated in Figure 6.3. With only two banks competing, the incumbent's price remains sufficiently higher than the entrant's for the latter to attract a considerable portion of the established market. Exit seems far less plausible in this model: with so few firms the collusive tendency common in switching cost models becomes dominant and the market tends to be shared - compare the entrant monopoly of new generation and incumbent monopoly of old generation results of Farrell and Shapiro's [1988] model. In the triopoly model the outcome is very different: E obtains only a small proportion of the L customers and is thus highly dependent on the N customers. If the business provided by these is not sufficiently great, remaining in the market is not worthwhile. Table 6.1 shows market shares resulting from first period pricing in the duopoly model and Table 6.2 provides, for comparison, the shares obtained when there are three firms competing in the entry period. It assumes that the two incumbents begin the game with equal market shares of the established customers.

Comparison of (6.20) and (6.21) shows that $h_{E1}(N^a) > h_{E1}(N^b)$ if $N^a > N^b$, as expected. Thus, at the end of period 1, we may specify the beliefs held by E conditional on demand and prices as: E assigns probability 1 to N^b if $p_{E1} h_{E1} \leq p_{E1} h_{E1}(N^b)$ and probability zero to N^b if $p_{E1} h_{E1} > p_{E1} h_{E1}(N^b)$. These beliefs are consistent Bayesian updating given N^a and N^b and the exit strategy is optimal given these beliefs.

---

[13] This facilitates comparison with the values in the appendix and suggests that the simplification of unitary δ made early in the appendix model does not affect our conclusions.

Figure 6.3: *First period duopoly market shares*

O_i = i's origin for N customers; O_E = E's origin for N customers. $O_i O_E$ = 1 = L measures the old customers. ——— : indicates market limits when $N = N^a$; — —: indicates market limits when $N = N^b$
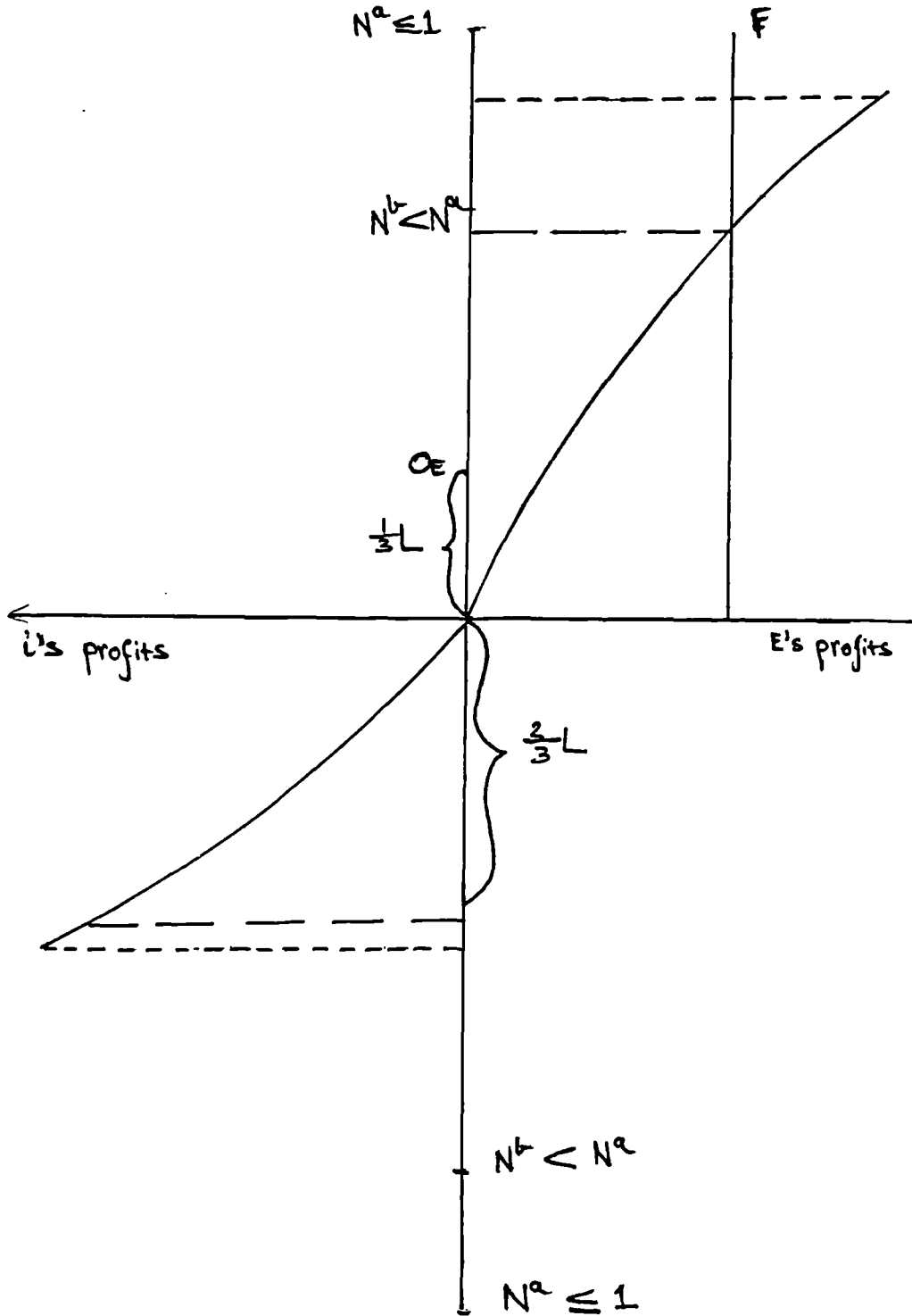
Table 6.1

Comparison of Stage 1 Duopoly Market *Shares*[*] (percentage)			
	N^a	N^b	L
$\rho = 0.8$			
incumbent, i	23.5	7.56	68.91
entrant, E	76.56	92.44	31.09
$\rho = 0.2$			
incumbent	5.0	27.48	67.18
entrant	95.0	72.52	32.82

Notes:

[*]: $\delta = 1$ in these calculations

Table indicates shares of the market segments in the column headings when the incumbent is a monopolist so that incumbent and entrant compete as a duopoly in period 1. Thus, in the event that $N = N^a$, with beliefs attaching 0.8 probability to that eventuality, the incumbent acquires 23.5% of the new customers (first entry).

Table 6.2

Comparison of Stage 1 Triopoly Market Shares (percentage)			
	N^a	N^b	L
$\rho = 0.8$			
incumbent i	28.2136	25.3939	48.1523
incumbent j	28.2136	25.3939	48.1523
entrant, E	43.5728	42.2124	3.6955
$\rho = 0.2$			
incumbent i	25.7779	26.3866	49.6621
incumbent j	25.7779	26.3866	49.6621
entrant, E	48.3034	47.2268	0.6758

Note: The table shows percentage share of each player in the market segments shown in the column headings when the incumbency is a duopoly so that entrant and incumbents compete as triopoly in period 1. We have assumed that i and j begin the game with equal shares, $l_{i0} = l_{j0} = 0.5$. Consequently, since their strategies are symmetric, their market shares are equal.

Stage 2 profits are obtained by substituting h_{k1} ($k = i,E$) into profit functions of the form in (6.10). Thus when E is in ($N = N^a$):

$$\Pi_{E2}^d = (27-4\delta\rho)^{-2}[(18-3\rho-2\delta\rho)N^a - 3(1-\rho)N^b + 2(6-\delta\rho)L]^2 \tag{6.23}$$

$$\Pi_{i2}^d = (27-4\delta\rho)^{-2}[(15-2\delta\rho)L + (9+3\rho-2\delta\rho)N^a + 3(1-\rho)N^b]^2 \tag{6.24}$$

As the rational expectations nature of the equilibrium suggests, we check that exit is indeed consistent with losses when $N = N^b$. While F has not been specified, if $\Pi_{E2}^d(N^a) > \Pi_{E2}^d(N^b)$, there will be values of F for which (6.3) is satisfied. E's second period profits when $N = N^b$ are given by (6.10) with h_{E1}^d (N^b) from (6.21) and h_{i1}^d from (6.22) substituted:

$$\Pi_{E2}^d(N^b) = [3(27-4\delta\rho)]^{-2}[(45+9\rho-8\delta\rho)N^b +6(6-\delta\rho)L -\rho(9-2\delta)N^a]^2 \tag{6.25}$$

and comparison of (6.23) and (6.25) shows that $N^a > N^b$ is required for $\Pi_{E2}^d(N^a) > \Pi_{E2}^d(N^b)$.

We also need to check that E's entry would be rational. That is, that the expected profit from entry is nonnegative: expected net revenue exceeds F. For this we require that expected demand in period 1 ($E\,h_{E1}$) and expected profits in period 2 exceed realized h_{E1} and Π_{E2}, respectively, evaluated at $N = N^b$. Expected period 1 demand is:

$$Eh_{E1} = \rho N^a + (1-\rho)N^b + p_{i1}-p_{E1}$$
$$= (27-4\delta\rho)^{-1}[\rho(18+2\delta-4\delta\rho)N^a +(1-\rho)(18-4\delta\rho)N^b +(9-2\delta\rho)L] \tag{6.26}$$

Again, the only requirement for $Eh_{E1}^d > h_{E1}^d$ (N^b) is $N^a > N^b$.

E's expected period 2 profits are $\rho\Pi_{E2}^d(N^a)+(1-\rho)0$ which is just (6.23) times ρ. They exceed realized profits at $N = N^b$ when:

$$[3\sqrt{\rho}(18-3\rho-2\delta\rho)+9-2\delta]N^a > [45+9\rho-8\delta\rho+9\sqrt{\rho}(1-\rho)]N^b \tag{6.27}$$
$$+ 6(6-\delta\rho)(1-\sqrt{\rho})L$$

Clearly, this condition is more easily satisfied for high values of ρ. Setting $\delta = 1$, when $\rho = 0.8$, satisfaction of (6.27) implies $N^a \geq 1.0638N^b + 0.0739$, and, taking N^a at its maximum value of unity, N^b must not exceed 0.8706 in order that (6.27) be satisfied. The permissible range for N^b is more restricted for low ρ. For example, a similar comparison when $\rho = 0.2$, requires N^b below 0.2358.

Calculation of market shares in the second period when E has remained in ($N = N^a$) emphasizes the collusive tendencies of the market when there is a monopoly incumbent. Second period duopoly demands when there is successful entry ($N = N^a$) are given for E and i by (6.28) and (6.29), respectively,

$$h_{E2}^d = (27 - 4\delta\rho)^{-1} [(18 - 3\rho - 2\delta\rho)N^a - 3(1-\rho)N^b + 2(6 - \delta\rho)L] \qquad (6.28)$$

$$h_{i2}^d = (27 - 4\delta\rho)^{-1} [(15 - 2\delta\rho)L + (9 + 3\rho - 2\delta\rho)N^a + 3(1-\rho)N^b] \qquad (6.29)$$

The calculation of market shares for these demands indicates that if $N = N^a$, so that E remains in, he obtains almost half the established customers, while keeping about 5/8 of the new. Given the separability of the functions, the second period share in L is invariant to the value of N so exit because of failure to obtain a sufficiently high number of the high value customers is unconvincing. In second period *triopoly*, however, he obtains only 20% of the more valuable customers while the incumbents capture about 30% [14] of the new customers (see Table A6). Exit therefore seems sensible in this framework only for more than two firms - which is not an unattractive result from an empirical point of view since we may otherwise be called upon to explain the presence of several banks in all the markets being considered.

6.5: Discussion and Conclusion

Specifying current demand by a rather general functional form with market share and price differences as arguments, Farrell showed that in the second period of a market with switching costs a large firm charges a higher price but not one that is so high as to erode its market lead. Using a linear version of this demand function we have expanded the model to consider the effects of first period pricing, the entry of new customers of uncertain value and the entry of new competitors. As in Farrell's model, market share has the status of a state variable which determines the strategic variable, price. The requirement that a firm enter, possibly to experience losses, in order to ascertain the value of new business is intended to capture an informational

---

[14] The ρ value used makes little difference to these proportions since, although influenced by the uncertainty of stage 1 prices, stage 2 prices are set when N's value has been learnt.

feature which appears more important for banks. They cannot, for example, conduct surveys and research to learn market potential, as can commodity sellers but must experience their customers[15]. Losses experienced in the first period, if staying in proves infeasible, are therefore the cost of information acquisition.

Incumbent banks are at an advantage in that they have a known customer base and no fixed costs. The entrants are acting on beliefs alone and we have restricted the information on which they condition their posteriors to their own observation of demand. It could be advantageous for an incumbent to attempt to influence either potential entrants' priors or posteriors since the incumbents stand to lose by entry, whether or not the entrant eventually exits. Post-entry attempts will give another equilibrium(a). We can informally consider what such an equilibrium would look like. By lowering price the incumbent could reduce the clientele available to an entrant and thus try to persuade him that $N = N^b$. Of course, the resulting price observed by the entrant would be inconsistent with the price expected given his priors and the entrant would not be fooled; this is not a convincing equilibrium since if $N = N^a$, the incumbent will have lost on price without having gained by exit. It could also be consistent with our story to assume that incumbents have better information about N and/or an entrant has less information about L. (We have assumed that the entrant knows the value of L customers, presumably because it is able to infer this pre-entry from its knowledge of incumbents' payoff functions). With such asymmetric information between competitors, we will be in the world of Roberts' [1986] model where the incumbent's privileged information about the strength of market demand produces predatory pricing to try and induce exit. Pre-entry strategic pricing for deterrence, of the type analyzed by Klemperer, op.cit., also seems a possible equilibrium response to threatened entry. With uncertainty about market demand, low price would both signal low business volume and build customer base. There are thus two counts on which entry may be dissuaded with lower prices.

---

[15] Thus the established banks are also uncertain of the value of new business.

Comparison of the monopoly incumbency and duopoly incumbency models suggests that in markets where loyalty/switching costs are important, the number of firms may be important to the equilibrium found. Given a market of fixed size or growing only slowly, (and switching costs are most significant in such markets) the number of firms determine the market share (customer base) which in turn determines strategic stance[16].

Consider the model in our empirical setting. There is some, probably regional, exogenous event which banks expect to expand the demand for their services. If their beliefs about prospects are sufficiently optimistic, multinationals enter the market to take advantage of the new business. They probably have a larger incentive to enter a market when their rivals are doing so - especially a market with switching costs where they are easily pre-empted. If they find it to be sufficiently lucrative they will stay in the market. But if the new business is insubstantial, and banks fail to acquire a sufficient share of the old, they will exit having acquired costly information about regional prospects[17]. Of course, if they enter several regional markets simultaneously they may well find one where staying is profitable. Having entered a market, the fewer the incumbent firms, the better their chances of staying in since larger banks are less aggressive in the face of entry. The cost of information acquisition therefore rises with the number of incumbent firms and we therefore expect the rate of entry to decrease over time. In an apparently more competitive market (one with more firms) entry proves to be harder - a collusive oligopoly is essentially its own protection here, provided it does not collude during the entry period. Even if there are only prudential restrictions on the issue of bank licences, a number of banks, among whom long familiarity may help tacit collusion, will not be seriously challenged by outsiders seeking profitable opportunities. The result is consistent with observations in Caribbean markets and suggests that one should not necessarily expect financial liberalization to encourage competitive entry.

---

[16] We may also recall Scotchmer's [1986] model showing there is no stationary Nash equilibrium for more than two firms where there are marginal customers with zero switching costs. Any firm with less than half the market is torn between increasing price and losing customers, and decreasing price and attracting more sales from both his own clientele and the marginal customers of other firms: demand becomes more elastic at lower prices.

[17] Domestic entrants will have acquired information about new areas of activity.

Appendix A6: The Exit Model with Triopoly in Stage 1

Equilibrium Behaviour

Here we consider the same problem as in Section 6.4 but assume that there is a duopoly incumbent. Normalizing the volume of established business, $L = \sum_{k=1}^{k=K} l_k$, where l_k is the market share of bank k, at unity, and denoting the volume of new business by N, any bank k acquires the following new customers in a period in which new customers enter the market:

$$\left[\frac{1-l_k}{K-1}\right] N$$

where K is the number of banks and division by $(K-1)$ ensures that market shares sum to N. In period t in which new customers enter, a bank on the market, say i, with rivals k and j has current demand:

$$h_{it} = l_{it-1} + (K-1)^{-1}(1-l_{it-1})N + (p_{jt}-p_{it}) + (p_{kt}-p_{it}) \tag{A6.1}$$

where $l_{it-1} = h_{it-1}$ is bank i's share of established customers, its demand resulting from pricing in the previous period. Total industry demand is assumed inelastic. For a single bank current market demand only exceeds inherited market share and the initial acquisition of new customers if $p_{jt} - p_{it} > 0$ and/or $p_{kt} - p_{it} > 0$. That is, the net price of bank i is below those of other banks. Demand in any period depends on both current and previous prices. With this function, when considering more than two firms, we obtain the rather unconvincing implication that, for any bank to attract more customers it need only price below the average price of other banks on the market. That is, from (A6.1)

$$h_{it} > h_{jt} \quad if \quad p_{it} < \frac{(p_{jt}+p_{kt})}{2}$$

In the usual interpretation, symmetry of the firms justifies this. Here firms differ because of their market share. But a similar justification is available if we treat long-established banks as having similar market shares, as we would expect. We will therefore be assuming that $l_{i0} = l_{j0} = 0.5$, although for the purpose of keeping track of the effect of market shares on price, we maintain the ls explicitly.

For a bank entering a previously duopolistic market in stage 1 at the same time as the new customers, expected stage 1 demand is given by

$$h_{E1} = \frac{1}{2}N^e + p_{j1} - p_{E1} + p_{i1} - p_{E1} \tag{A6.2}$$

and for an incumbent bank, i, by:

$$h_{i1} = l_{i0} + \frac{1}{2}(1 - l_{i0})N^e + p_{j1} - p_{i1} + p_{E1} - p_{i1} \tag{A6.3}$$

where l_{i0} is bank i's share of established customers from the pre-entry stage 0. In (A6.1) $K = 3$, and the entering bank has no established customers $(l_{E0} = 0)$.

In equilibrium in stage 2 the players will have learnt the magnitude of N from their own demands. They also know whether or not E has exited. Banks therefore solve a one period problem. If E has exited we assume that its customers distribute themselves randomly among the remaining banks in the market.

If $N = N^a$, E remains in the market and each bank i chooses p_{i2} to maximize:

$$\Pi_{i2} = p_{i2}h_{i2} = p_{i2}[h_{i1}(N^a) + \bar{p}_{j2} - p_{i2} + \bar{p}_{E2} - p_{i2}] \tag{A6.4}$$

where i and j denote the incumbents. $\bar{p}_{j2}$ is the price i conjectures that j will set. The maximand is symmetric for j and E. $h_{i1}(N^a)$ is the demand resulting from pricing in Stage 1 when $N = N^a$. This gives best response functions:

$$p_{i2} = \frac{1}{4}[h_{i1}(N^a) + \bar{p}_{j2} - \bar{p}_{E2}] \tag{A6.5}$$

In triopoly (denoted T) equilibrium, i's price is

$$p_{i2}^T = \frac{1}{10}[3h_{i1}(N^a) + h_{E1}(N^a) + h_{j1}(N^a)] \tag{A6.6}$$

As discussed with reference to equation (6.8), price depends positively on the market shares of all banks, with greater weight given to own share. Substituting prices of the form in (A6.6) for i, j and E into the demand function in (A6.4) evaluated at equilibrium prices, realized demand for i is

$$h_{i2}^T = \frac{1}{5}[3h_{i1}(N^a) + h_{E1}(N^a) + h_{j1}(N^a)] \tag{A6.7}$$

and i will have increased its market share $(h_{i2} > h_{i1})$ if its previous share is less than the average

of the other banks. Or, assuming, as discussed above, that established banks begin with similar shares, the entrant (an incumbent) increases its market share in period 2 if the business acquired in period 1 is less than that of the incumbents' (entrant's): because its inherited market share is less, its period 2 price is also lower since it has a greater incentive to acquire more customers. Multiplying (A6.6) and (A6.7), profits are given by

$$\Pi_{i2}^{T} = \frac{1}{50}[3h_{i1}(N^{a}) + h_{E1}(N^{a}) + h_{j1}(N^{a})]^{2} \tag{A6.8}$$

Substituting for the period 1 demands from (A6.2) and (A6.3) the triopoly profits for i in period 2 can be expressed in terms of stage 1 prices:

$$\Pi_{i2}^{T} = \frac{1}{50}[3l_{i0} + l_{j0} + \frac{1}{2}N^{a}(5 - 3l_{i0} - l_{j0}) + 2\bar{p}_{j1} + 2\bar{p}_{E1} - 4p_{i1}]^{2} \tag{A6.9}$$

If E has exited $N = N^{b}$, then i (and j symmetrically) maximize

$$\Pi_{i2} = p_{i2}[h_{i1}(N^{b}) + \frac{1}{2}h_{E1}(N^{b}) + \bar{p}_{j2} - \bar{p}_{i2}] \tag{A6.10}$$

where it is assumed that E's abandoned customers open accounts at random with the remaining banks. The best reply functions are given by:

$$p_{i2} = \frac{1}{2}[h_{i1}(N^{b}) + \frac{1}{2}h_{E1}(N^{b}) + \bar{p}_{j2}] \tag{A6.11}$$

Given consistent price conjectures, solving (A6.11) for i and j simultaneously, equilibrium prices are given by

$$p_{i2}^{D} = \frac{1}{3}[2h_{i1}(N^{b}) + \frac{3}{2}h_{E1}(N^{b}) + h_{j1}(N^{b})] \tag{A6.12}$$

$$= h_{i2}^{D},$$

the resulting share of bank business obtained in stage 2. Superscript D denotes duopoly equilibrium variables for the Appendix model. Duopoly profits are therefore given by:

$$\Pi_{i2}^{D} = \frac{1}{9}\left[2h_{i1}(N^{b}) + \frac{3}{2}h_{E1}(N^{b}) + h_{j1}(N^{b})\right]^{2} \tag{A6.13}$$

Substituting in (A6.13) for h_{i1}, h_{E1} and h_{j1} from (A6.2) and (A6.3), given the realized value of N, stage 2 profits in terms of stage 1 prices are:

$$\Pi_{i2}^{D} = \frac{1}{9}\left[2l_{i0} + l_{j0} + \frac{1}{2}N^{b}(4.5 - 2l_{i0} - l_{j0}) + 1.5\bar{p}_{j1} - 1.5p_{i1}^{'}\right]^{2} \tag{A6.14}$$

Consider the first period. The incumbents know that E has entered and its exit rule. The distribution of N is common knowledge. Current price influences current demand as well as the volume of business that the banks will have in the second period. The entrant therefore maximizes expected discounted profits:

$$E\Pi_E + (1+\delta\rho)F = E\Pi_{E1} + \delta\rho\Pi_{E2}^T(N^a)$$

Substituting from (A6.2) and (A6.3) into Π_{E1} and Π_{E2}^T, this maximand becomes:

$$E\Pi_E = p_{E1}[\tfrac{1}{2}N^e + \bar{p}_{i1} + \bar{p}_{j1} - 2p_{E1}]$$

$$+ \tfrac{\delta\rho}{50}[\tfrac{1}{2}N^a(5 - l_{i0} - l_{j0}) + l_{i0} + l_{j0} + 2\bar{p}_{i1} + 2\bar{p}_{j1} - 4p_{E1}]^2 \qquad (A6.15)$$

where δ is the common discount factor. The incumbents also maximize expected discounted profits:

$$\Pi_i = E\Pi_{i1} + \delta\rho\Pi_{i2}^T(N^a) + \delta(1-\rho)\Pi_{i2}^D(N^b)$$

and, using (A6.9) and (A6.14)

$$\Pi_i = p_{i1}[l_{i0} + \tfrac{1}{2}\rho(1-l_{i0})N^a + \tfrac{1}{2}(1-\rho)(1-l_{i0})N^b + \bar{p}_{j1} + \bar{p}_{E1} - 2p_{i1}]$$

$$+ \tfrac{1}{9}\delta(1-\rho)[2l_{i0} + l_{j0} + \tfrac{1}{2}N^b(4.5 - 2l_{i0} - l_{j0}) - 1.5p_{i1} + 1.5\bar{p}_{j1}]^2$$

$$+ \tfrac{1}{50}\delta\rho[3l_{i0} + l_{j0} + \tfrac{1}{2}N^a(5 - 3l_{i0} - l_{j0}) + 2\bar{p}_{j1} + 2\bar{p}_{E1} - 4p_{i1}]^2 \qquad (A6.16)$$

The first order conditions give the best response functions

$$p_{E1} = (100 - 16\delta\rho)^{-1}\left[\tfrac{1}{2}\rho(25 - 20\delta + 48L)N^a + \tfrac{25}{2}(1-\rho)N^b - 48\rho L + (25 - 88\rho)(\bar{p}_{i1} + \bar{p}_{j1})\right] \qquad (A6.17)$$

$$p_{i1} = (200 - 25\delta - 78\rho)^{-1}\left\{\tfrac{2}{3}\left[75 - 50\delta + 14\delta\rho\right]l_{i0} - \tfrac{2}{3}\left[25\delta - 13\delta\rho\right]l_{j0}\right. \qquad (A6.18)$$

$$+ \rho\left[25(1-l_{i0}) - 20\delta + 48(3l_{i0} + l_{j0})\right]N^a$$

$$+ (1-\rho)\left[25(1-l_{i0}) - 37.5\delta + \tfrac{25\delta}{3}(2l_{i0} - l_{j0})\right]N^b$$

$$\left. + (50 - 25\delta + 98\rho)\bar{p}_{j1} + 2(25 - 88\rho)\bar{p}_{E1}\right\}$$

Solving (A6.17) and (A6.18) (with δ set equal to unity for ease of calculation) we obtain the first period equilibrium prices:

$$p_{E1} = [5(250 - 488\delta)]^{-1}\left\{\tfrac{1}{2}\rho\left[(125 - 24\rho) + (15 + 8\rho)L\right]N^a\right. \qquad (A6.19)$$

$$+ \tfrac{1}{2}(1-\rho)\left[2.5(200-32\rho)+\tfrac{10}{3}(25-8\rho)L \right] N^b -\rho(15+8\rho)L\bigg\}$$

$$p_{i1} = [\,5(200+2\rho)(150-16\rho)(250-48\rho)\,]^{-1} \tag{A6.20}$$

$$\left\{ \left[\tfrac{10}{3}(250-48\rho)[25^2(6-3.8\rho)+19\rho^2] - 2\rho(15+8\rho)(25-8\rho)(200+2\rho) \right] l_{i0} \right.$$

$$- \left[\tfrac{10}{3}(250-48\rho)[25^2(6-4.84\rho)-35\rho^2] + 2\rho(15+8\rho)(25-8\rho)(200+2\rho) \right] l_{j0}$$

$$+ \rho\left[(200+2\rho)[(125-24\rho)(75-8\rho) + (15+8\rho)(25-8\rho)L\,] + 50(125-24\rho)[(75-29\rho)l_{j0} - 5(87-5.08\rho)l_{i0}] \right] N^a$$

$$\left. + (1-\rho)(200+2\rho)\left[125(125-24\rho)+5(25-8\rho)(100-16\rho)+\tfrac{10}{3}[25(125-24\rho)+(25-8\rho)^2]L \right] N^b \right\}$$

In contrast to the second period prices (compare (A6.6)) we note that price is decreasing in rivals' business volume at the beginning of period 1 (negative coefficients on L and l_{j0} in (A6.19) and (A6.20), respectively). Whereas consideration of the single period problem illustrates only one aspect of the incentives in markets with switching costs - that of pricing high to exploit customer base - the forward-looking price reflects the incentive to price low in order to build a customer base for the next period. This tendency is more profitable, the higher the current share of rivals, and hence price is negatively related to rivals' share. On the other hand, a high starting share of old customers (high l_{i0}) reduces the starting share of new customers and hence has a negative impact on price when $N = N^a$, since lower price will attract new customers if E remains in the market. If E exits ($N = N^b$) incumbents have less incentive to attract N business since they will in any case obtain E's share.

It is easily seen that $p_{i1} \geq (\leq) p_{j1}$ as $l_{i0} \geq (\leq) l_{j0}$. However, a comparison between E's price and that of the incumbents is more difficult. The components of net price are always greater for the market segment in which a bank "specializes", e.g. the coefficient of L in p_{i1} exceeds its coefficient in p_{E1}. We would also expect the entrant to price lower overall given its need to attract reliable clientele, and this is indeed so in the single period problem. In a single period game, where N was known to have value n, equilibrium prices are

$$p_{E1} = \tfrac{1}{10}(2n + l_{j0} + l_{i0}) \tag{A6.21}$$

$$p_{i1} = \tfrac{1}{10}(3l_{i0} + l_{j0} + (2-l_{i0})n)$$

- 151 -

However, in the two stage game E's incentive to offer lower prices is reduced by the possibility of exit before it could gain the full returns on those prices. The differences between p_{i1} and p_{E1} are given by (with $L=1$, $l_{i0} = l_{j0} = 0.5$)

$$p_{i1} - p_{E1} = [3(200+2\rho)(150-16\rho)(250-48\rho)]^{-1}$$
$$\left\{ 2\rho(250-48\rho)(325+27\rho) + 60\rho(15+8\rho)(200+2\rho) \right.$$
$$- \rho(100+\rho)[60(15+8\rho) + 27(250-48\rho)]N^a$$
$$\left. - (1-\rho)25(100+\rho)(150-16\rho)N^b \right\}$$

(A6.22)

For the L segment p_{i1} exceeds p_{E1} and vice versa for N. Price differences for L and N^a are greater at high values of ρ since banks are attempting to differentiate themselves in anticipation of fiercer competition next period - prices on others' shares are lower, but price on own share is higher for greatest current benefit. But for N^b the excess of p_{E1} over p_{i1} is larger at low ρ since E then tries to make the most of current opportunity.

As a result of these prices, in period 1 banks obtain the following business which varies according to whether realized N is N^a or N^b:

$$h_{E1}(N^a) = [(100+\rho)(150-16\rho)(250-48\rho)]^{-1}$$
$$\left\{ (100+\rho)(75-8\rho)[(250-48\rho)-34\rho L]N^a \right.$$
$$- 10(1-\rho)(100+\rho)(75-8\rho)(5-\tfrac{10}{3}L)N^b$$
$$\left. + \rho\left[\tfrac{1}{3}(250-48\rho)(650+54\rho) + 40(15+8\rho)(100+\rho) \right]L \right\}$$

(A6.23)

$$h_{i1}(N^a) = [(100+\rho)(150-16\rho)(250-48\rho)]^{-1}$$
$$\left\{ (250-48\rho)[(100+\rho)(150-16\rho) - 25^2(6-4.1467\rho) - \rho^2]l_{i0} \right.$$
$$+ (250-48\rho)[25^2(6-4.4933\rho) - 17\rho^2]l_{j0}$$
$$- 20\rho(15+8\rho)(100+\rho)L$$
$$+ 2.5[(250-48\rho)[\rho(945-79.8\rho)l_{i0}+(3000-\rho(875-80.2\rho))l_{j0}] + 4\rho(15+8\rho)(100+\rho)L]N^a$$
$$\left. + 5(1-\rho)(100+\rho)(75-8\rho)(5-\tfrac{10}{3}L)N^b \right\}$$

(A6.24)

Summing over these demands for all three banks gives total market demand

$$\sum_{k=i,j,E} h_{k1} = L + N^a$$

so that by specifying values for ρ we may obtain relative shares. Although E's price for N exceeds i's, E keeps most of its new customers, whatever the value of N and ρ. And it obtains only a small proportion of the L customers - 0.68% when $\rho = 0.2$ and 3.7% when $\rho = 0.8$, see Table 6.2. Since the excess of p_{i1} is larger for high ρ, the proportion obtained by E is higher in the latter case. E is thus highly dependent on the N customers. If the business provided by these is not sufficiently great, remaining in the market is not worthwhile.

When $N = N^b$, realized demands at the end of period 1 are given by:

$$h_{E1}(N^b) = [(100+\rho)(150-16\rho)(250-48\rho)]^{-1} \tag{A6.25}$$
$$\left\{ (100+\rho)(75-8\rho)\left[(250-48\rho) - 10(1-\rho)(5-\tfrac{10}{3}L) \right] N^b \right.$$
$$- 34\rho(100+\rho)(75-8\rho)N^a$$
$$\left. + \rho\left[(250-48\rho)(216.6667+18\rho) + 40(100+\rho)(15+8\rho) \right] \right\}$$

$$h_{i1}(N^b) = [(100+\rho)(150-16\rho)(250-48\rho)]^{-1} \tag{A6.26}$$
$$\left\{ \left[(100+\rho)[(150-16\rho)(250-48\rho) - 20\rho(15+8\rho)] - (250-48\rho)[25^2(6-4.1467\rho)+\rho^2] \right] l_{i0} \right.$$
$$+ \left[(250-48\rho)[25^2(6-4.4933\rho) - 17\rho^2] - 20\rho(100+\rho)(15+8\rho) \right] l_{j0}$$
$$+ (100+\rho)(75-8\rho)\left[(1-l_{i0})(250-48\rho) + 5(1-\rho)(5-\tfrac{10}{3}L) \right] N^b$$
$$\left. + 5\rho\left[2L(100+\rho)(15+8\rho) + (125-24\rho)[(945-79.8\rho)l_{i0} - (585-83.4\rho)l_{j0}] \right] N^a \right\}$$

Comparison of (A6.23) and (A6.25), and of (A6.24) and (A6.26) shows that $h_{k1}(N^a) > h_{k1}(N_b)$ for $k = i,j,E$ if $N^a > N^b$, as expected. Thus, at the end of period 1, we may specify the beliefs held by E conditional on demand and prices as: E assigns probability 1 to N^b if

$$p_{E1}h_{E1} \leq p_{E1}h_{E1}(N^b)$$

and probability zero to N^b if

$$p_{E1}h_{E1} > p_{E1}h_{E1}(N^b)$$

These beliefs are consistent Bayesian updating given N^a and N^b and the exit strategy is optimal given these beliefs.

Stage 2 profits are obtained by substituting h_{k1} ($k = i,j,E$) into the profit functions of the forms in (A6.8) and (A6.12). Thus when E is in ($N = N^a$):

$$\Pi_{E2}^T + F = \tfrac{1}{50}[(100+\rho)(150-16\rho)(250-48\rho)]^{-2} \tag{A6.27}$$

$$\left\{ (100+\rho)(75-8\rho)\left[(250-48\rho)(5-L) - 68\rho L \right] N^a \right.$$

$$- 20(1-\rho)(100+\rho)(75-8\rho)(5-\tfrac{10}{3}L)N^b$$

$$\left. + \left[(250-48\rho)[(150-16\rho)(100+\rho) + 433.3125\rho + 36\rho^2] + 80\rho(15+8\rho)(100+\rho)\right] L \right\}^2$$

$$\Pi_{i2}^T = \tfrac{1}{50}[(100+\rho)(150-16\rho)(250-48\rho)]^{-2} \tag{A6.28}$$

$$\left\{ \left[(250-48\rho)[3(100+\rho)(150-16\rho) - (25^2(12-8.2934\rho)-\rho^2)]\right] l_{i0} \right.$$

$$+ \left[(250-48\rho)[(100+\rho)(150-16\rho) + (25^2(12-8.9866\rho)-34\rho^2)]\right] l_{j0}$$

$$- 40\rho(100+\rho)(15+8\rho)L$$

$$+ \left[6\rho(100+\rho)(525+8\rho) + (250-48\rho)[5(100+\rho)(75-8\rho) \right.$$

$$\left. - 3(25^2(12-6.2\rho)-274\rho^2)l_{i0} - (25^2(12+8.2\rho)+826\rho^2)l_{j0}]\right] N^a$$

$$\left. + (1-\rho)\tfrac{50}{3}(100+\rho)(75-8\rho)N^b \right\}^2$$

With E's exit ($N = N^b$), the incumbents' profits are given by (A6.29):

$$\Pi_{i2}^D = [3(100+\rho)(150-16\rho)]^{-2} \tag{A6.29}$$

$$\left\{ \left[2(100+\rho)(150-16\rho) - (25^2(6-4.32\rho)-8\rho^2)\right] l_{i0} \right.$$

$$+ \left[(100+\rho)(150-16\rho) + (25^2(6-4.32\rho)-8\rho^2)\right] l_{j0}$$

$$+ (100+\rho)(75-8\rho)(3.5-l_{i0})N^b$$

$$\left. + \rho\left[2.5[(1305-76.2\rho)l_{i0}-(225-87\rho)l_{j0}] - 13.5(100+\rho)\right] N^a \right\}$$

As the rational expectations nature of the equilibrium suggests, we check that exit is indeed consistent when $N = N^b$. While F has not been specified, if $\Pi_{E2}^T(N^a) > \Pi_{E2}^T(N^b)$, there will be

values of F for which

$$\Pi_{E2}^{T}(N^{b}) \leq F < \Pi_{E2}^{T}(N^{a})$$

E's triopoly profits when $N = N^{b}$ are given by an equation of the form (A6.8) with appropriate substitutions of h_{k1}s:

$$\Pi_{E2}^{T} + F = \frac{1}{50}[(100+\rho)(150-16\rho)(250-48\rho)]^{-2} \qquad (A6.30)$$

$$\left\{ (100+\rho)(75-8\rho)\left[(250-48\rho)(5-L) - 20(1-\rho)(5-\frac{10}{3}L) \right] N^{b} \right.$$

$$- 68\rho L (100+\rho)(75-8\rho) N^{a}$$

$$+ \left[(250-48\rho)[(150-16\rho)(100+\rho) + 433.3125\rho + 36\rho^{2}] + 80\rho(15+8\rho)(100+\rho) \right] L \right\}$$

and $\Pi_{E2}^{T}(N^{a}) > \Pi_{E2}^{T}(N^{b})$ if

$$N^{a} (100+\rho)[(75-8\rho)[(250-48\rho)(5-L) - 68\rho L] + 68\rho L (75-8\rho)] > N^{b} (100+\rho)(75-8\rho)(250-48\rho)(5-L)$$

which clearly holds given $N^{a} > N^{b}$

We also need to check that E's entry would be rational. That is, that the expected profit from entry is nonnegative: expected net revenue exceeds F. For this we require that expected demand in period 1 ($E\, h_{E1}$) (recall that p_{E1} is invariant to realized type) and expected profits in period 2 exceed realized h_{E1} and Π_{E2}, respectively, evaluated at $N = N^{b}$. Expected period 1 demand is:

$$E\, h_{E1} = [(100+\rho)(150-16\rho)(250-48\rho)]^{-1} \qquad (A6.31)$$

$$\left\{ \rho(100+\rho)(75-8\rho)\left[(250-48\rho) - 34L \right] N^{a} \right.$$

$$+ (1-\rho)(100+\rho)\left[100(150-48\rho) - 8\rho(50-48\rho) + \frac{100}{3}(75-8\rho)L \right] N^{b}$$

$$+ \rho\left[(250-48\rho)(216.6667+18\rho) + 40(15+8\rho)(100+\rho) \right] L \right\}$$

and we find $E\, h_{E1} > h_{E1}(N^{b})$ if $N^{a} > N^{b}$

E's expected period 2 profits are given by $\rho \Pi_{E2}^{T}(N^{a}) + (1-\rho)0$ Expected profits in period 2 will exceed realized profits for E when $N = N^{b}$ if

$$4(100+\rho)(75-8\rho)[\sqrt{\rho}(250-48\rho) + 17\rho(1-\sqrt{\rho})]N^a \quad > \quad \text{(A6.32)}$$

$$(100+\rho)(75-8\rho)[4(250-48\rho) - \tfrac{100}{3}(1-\rho)(1-\sqrt{\rho})]N^b$$

$$+ (250-48\rho)[(150-16\rho)(100+\rho) + 433.3125\rho + 36\rho^2](1-\sqrt{\rho}) + 80\rho(15+8\rho)(100+\rho)(1-\sqrt{\rho})$$

This condition is more easily satisfied for high values of ρ. For example, taking $\rho = 0.8$, inequality (A6.32) is satisfied when $N^a > 1.1087 N^b + 0.0623$. Setting N^a at its maximum value of unity, this allows N^b to take any value below 0.8458. But when $\rho = 0.2$, for example, (A6.32) is satisfied for $N^a > 2.1639 N^b + 0.6156$ and, for $N^a = 1$, N^b cannot exceed 0.1776. Thus, if the probability attached to a high value of new customer business is low, it will be irrational for the entrant to seek out information in the market.

Even if E were to remain in the market in the final period (when $N = N^a$), it obtains a minority share of L customers. Realized second period demands in this case are given by:

$$h^T_{E2}(N^a) = [5(100+\rho)(150-16\rho)(250-48\rho)]^{-1} \quad \text{(A6.33)}$$

$$\left\{ (100+\rho)(75-8\rho)\left[4(250-48\rho) - 68\rho \right] N^a \right.$$

$$- (1-\rho)\tfrac{100}{3}(100+\rho)(75-8\rho)N^b$$

$$\left. + \left[(250-48\rho)[(150-16\rho)(100+\rho) + 433.3125\rho + 36\rho^2] + 80\rho(15+8\rho)(100+\rho) \right] \right\}$$

$$h^T_{i2}(N^a) = [5(100+\rho)(150-16\rho)(250-48\rho)]^{-1} \quad \text{(A6.34)}$$

$$\left\{ \left[(250-48\rho)[3(100+\rho)(150-16\rho) - 25^2(12-8.2934\rho) + \rho^2] - 40\rho(100+\rho)(15+\rho) \right] l_{i0} \right.$$

$$+ \left[(250-48\rho)[(100+\rho)(150-16\rho) - 25^2(12-8.9866\rho) + 34\rho^2] - 40\rho(100+\rho)(15+8\rho) \right] l_{j0}$$

$$+ \left[6\rho(100+\rho)(525+8\rho) + (250-48\rho)[5(100+\rho)(75-8\rho) \right.$$

$$\left. - 3(25^2(12-6.2\rho) - 274\rho^2) l_{i0} - (25^2(12-8.2\rho) + 826\rho^2) l_{j0} \right] N^a$$

$$\left. + (1-\rho)\tfrac{50}{3}(100+\rho)(75-8\rho)N^b \right\}$$

Equations (A6.33) and (A6.34) are obtained by substituting the realized values of h_{k1}, as appropriate, into (A6.7)). Setting $l_{i0} = l_{j0} = 0.5$, when $\rho = 0.2$, E obtains 20.2703% of the L customers and maintains 39.4343% of the N customers, whereas, when $\rho = 0.8$, E obtains 21.4781% of L and keeps 31.2854% of N - see Table A6.

Table A6

Comparison of Stage 2 Triopoly Market Shares (percentage)			
	N^a	$\overset{\bullet}{l}_{i0}$	L
$\rho = 0.8$			
incumbent i	31.2854	54.6798	39.2624
incumbent j	31.2854	23.8449	39.2624
entrant, E	37.4291		21.4784
$\rho = 0.2$			
incumbent i	30.2829	51.2044	39.8988
incumbent j	30.2829	28.5931	39.8988
entrant, E	39.4343		20.2703

Notes: The table shows percentage share of each player in the market segments shown in the column headings. We have assumed that i and j begin the game with equal shares, $l_{i0} = l_{j0} = 0.5L = 0.5$. Consequently, since their strategies are symmetric, their market shares are equal.

* : The entries under l_{i0} indicate that, for example, with beliefs such that $\rho = 0.8$, i maintains 54.6798% of the business with which he started if E does not exit, j acquires 23.8449% of i's initial customers and E acquires 21.4784% from each of them.

Chapter 7: An Informational Role for Banks in the Credit Market: a Screening Model

7.1: Introduction

In this chapter we simplify and adapt Milde and Riley's [1988] model of sorting by loan size in order to examine the welfare implications of bank screening. Our adaptation accomplishes the following. First, it shows explicitly how banks' information acquisition through screening can affect allocation: loans are allocated according to risk category and not expected return. Second, we find a separating sequential equilibrium in which lower risk/lower productivity firms signal their type by accepting a lower loan size. Credit market screening models usually find a reactive equilibrium or, given assumptions about borrower responses, a Nash equilibrium. Reactive equilibrium is unattractive because it requires a change in the basic Nash assumption about strategic behaviour; and the Nash equilibrium is inappropriate to the implicit two-stage formulation of screening games. Our demonstration that the same assumptions that give Nash equilibrium are sufficient for its sequential refinement is therefore of interest. Third, it provides simple demand functions which we will use in Chapter 8 to show that incentive compatibility may mean that the envelope theorem no longer applies to screened borrowers.

Consideration of how asymmetric information in the bank credit market affects allocation and efficiency has centred on the Stiglitz-Weiss analysis of adverse selection and moral hazard sources of credit rationing. More recent analyses argue that banks can avoid the credit rationing response to adverse selection by offering contracts among which customers choose, that choice revealing their private information. These contracts may variously feature collateral, loan size, loan receipt probability etc. as screening instruments. Moral hazard effects may also be avoided by appropriate contract-fixed incentives.

Screening models offer theoretical justification of the original McKinnon-Shaw liberalization hypotheses: in lagging economies with imperfect information, financial intermediaries can act as a mechanism to elicit information which facilitates credit allocations. However, three further factors should be taken into account. First, theoretical models do not always predict a

screening outcome. Second, even with screening, perfect sorting may not occur. Third, there is an expanding theoretical literature which indicates that, where markets are incomplete, or information asymmetric, the equilibrium resulting from decentralized economic activity is not constrained Pareto efficient.

In Chapter 8 the loan demands and rates obtained here are incorporated into the general equilibrium framework used by Greenwald and Stiglitz [1986] to examine the Pareto efficiency of economies with incomplete markets and information constraints. We can thus examine the third factor mentioned above.

In Section 7.2 we survey models of potential-borrower discrimination by banks - whether through rationing or the use of screening devices. Section 7.3 discusses Caribbean credit markets in the light of the analyses. We suggest that it is difficult to unambiguously explain observed outcomes on a theoretical basis alone. Section 7.4 describes the model which assumes perfect sorting: banks go as far as they are able in eliciting incentive-compatible information from potential customers.

7.2: Asymmetric Information in the Credit Market: A Survey

The literature on asymmetric information in the the credit market examines how banks make their loan decisions when they face heterogeneous borrowers and lack direct access to reliable information on borrower characteristics or choices. It suggests that, unless banks have sufficient instruments (debt contract variables) relative to borrower characteristics/decisions at their disposal, and borrowers' relative preferences are appropriately structured, banks' optimal reactions to asymmetric information may entail the rationing of credit. More recent papers discuss the strategic and structural features which influence the marketability of these instruments.

Informational asymmetries in credit markets occur because potential borrowers have private information about their risk characteristics, and can adopt loan uses, which are not observed by banks. Following project implementation they also have more information about their ability to repay. Even if outcomes are observable, when there is a stochastic component banks cannot perfectly infer actions from the outcomes themselves.[1] To the extent that banks are able to obtain

---

[1] Perfect inference would allow contingent contracts and hence resolve moral hazard problems.

reliable information by investigations and monitoring, the problems of asymmetric information are attenuated. But a reasonable assumption is that directly obtaining all necessary information for all potential borrowers would be prohibitively costly. In addition to learning and direct information-gathering, the literature offers two methods by which lenders cope with asymmetric information: they may ration credit to deter the worst risks, or offer contracts to differentiate among risk classes and provide incentives.

Two major categories of credit rationing have been distinguished[2]. Rationing of loan size occurs where, at the quoted rate of interest, borrowers receive a smaller loan than demanded even though willing to pay a higher rate of interest. This is the definition used by Jaffee and Russell [1976], the first authors to analyse the effects of asymmetric information on the loan market.[3] The second category of credit rationing occurs where some borrowers are unable to obtain credit even though they would be willing to pay a higher rate of interest in order to do so. Stiglitz and Weiss [1981, 1983a] distinguish three possibilities in this second category. Borrowers denied loans may be observationally indistinguishable from those who receive; the borrowers denied loans may be identifiable groups; or loans may be denied in later periods as a result of earlier default, even though the later projects are better and the borrowers are willing to pay higher rates of interest.

Neither category of credit rationing would occur under common information even with uncertainty. Debt contracts would specify the (verifiable) purposes for which loan proceeds could be used and repayment would be made contingent on the (again verifiable) outcome. Contract terms (the rate of interest) would be set such that borrowers are indifferent between receiving or not receiving a loan, or larger loan, and to compensate the bank for risk, and the degree of competition among banks would determine how surplus is shared between bank and borrower.

The alternative to rationing are loan contracts which induce borrowers to reveal information about their risk type by selecting a contract, or give them the incentive to undertake the projects considered desirable by the bank. This is most easily understood in the context of the general

---

[2] They have similar sources. But rationing by loan size cannot occur where it is assumed that loan requirements are fixed.

[3] Although we note that Akerlof's pioneering article [1970] discussed imperfect information in credit markets in India.

theory of contracting between principal and agent, relevant features of which are described below.

The background of principal-agent contracting games

Models examining creditor strategies under asymmetric information can be treated as applications of the economics of information to the market for credit. Most recent models assume hidden information, or adverse selection, where the borrower is the agent or informed and the bank the principal or uninformed. With hidden information the components of the principal's contract offers will be aimed at eliciting information and differentiating among the agents. It is usually assumed that there is a one-dimensional distribution of agents who differ in quality (or risk). These differences are assumed to imply that agents have different marginal rates of substitution (MRS) between contract components. Where the MRS differ sufficiently among individuals, and it is profitable for the principal to have agents choose different treatment, i.e. to self-select, the market will be characterised by a separating equilibrium where each type of agent receives a different contract[4]. It may not however be possible (or desirable from the viewpoint of the principal) to induce perfect self-selection. Different types are then pooled together at a single contract. It is in this case that credit rationing usually occurs. If a bank's expected return from a loan contract is maximized at a rate of interest below that which equates the supply and demand of borrowers accepting that contract, some of those borrowers may be rationed out of the market, or may receive a smaller loan than they would prefer at the given rate of interest. Rationing is then one means by which banks may deal with the effects that unobserved risk differences have on their gross expected return, and the ability to offer sorting contracts is another.

In a game theory context, the description above assumes that it is the uninformed (the banks) who move first. All (published) credit rationing models make this assumption, i.e. they are screening rather than signalling (where the informed move first) models, in Stiglitz-Weiss's [1983b] terminology. This labelling is useful because it aids in distinguishing between empirical

---

4 Feasible separation requires that the difference in MRS can be ranked by type or, equivalently, that the screening variable (or signal) have a lower marginal cost for the bank-favoured type; that is, indifference curves must satisfy the single crossing property.

counterparts and draws attention to the fact that the equilibrium outcome expected should depend on the game structure or order of moves posited.

Models of asymmetric information markets tend to find too many or no Nash equilibria[5]. The existence and location of equilibrium may appear arcane concerns in the explanation of empirical phenomena. However, if a theoretical analysis is unable to predict a plausible equilibrium under the widely used Nash assumption, it suggests that something is wrong with the model's view of reality. Either the behavioural postulate is an incorrect description of decision-making in the markets modelled, or the market assumed does not exist; that is, there is market failure in the sense that there is a commodity[6] for which decentralized activity has failed to fix a price.

Various approaches have been used in attempts to pin down an equilibrium. As will become clear, the difficulty basically arises from the nature of competitive behaviour. Earlier answers changed the equilibrium behaviour of the competitors or the assumptions regarding the response of their customers. More recently, analysts have used the explicit approaches to strategic competitive behaviour provided by game theory. In the following, we focus on the credit market-relevant case of screening but begin by briefly considering signalling.

If it is assumed that the informed move first (signalling) with the uninformed reacting passively to their offers[7] multiple sequential[8] equilibria are obtained. Essentially, this is because the optimal reaction of the uninformed to the signal[9] will vary with their assessment of the information conveyed by the signal. As a result the informed chose a signal that depends on his anticipations of the reaction. Numerous off-the-equilibrium-path beliefs can be formulated for the uninformed, so it is difficult to rule out equilibria. By placing restrictions on the out-of-equilibrium

---

[5] Unless appropriate assumptions, which we describe below, are made.

[6] In Debreu's [1959] sense.

[7] Such as firms quoting a price to consumers uninformed about quality and individuals choosing education prior to seeking a job.

[8] A strategy combination is a sequential equilibrium if there are consistent beliefs such that each player's strategy prescribes at every information set a choice which is optimal with respect to those beliefs (Kreps and Wilson [1982]).

[9] Where the signal is the offer made by the informed.

beliefs of the uninformed, Cho and Kreps [1987] show in a three-stage game that an intuitive criterion can be used to eliminate unreasonable sequential equilibria, yielding a separating equilibrium.

Where the uninformed move first (screening)[10] "nonexistence (of Nash equilibria in pure strategies) is generic" (Riley, [1985], p.959) in a wide class of models, as in the original Rothschild and Stiglitz [1976] and Wilson [1977] insurance market papers[11]. It is the myopic nature of the noncooperative (Nash) reactions of the uninformed principals that eliminates equilibria. One answer has therefore been to adopt equilibrium concepts which abandon the Nash notion that players take others' strategies as given. The most widely used[12] is Riley's [1979] reactive equilibrium. It requires that any initial defector from an equilibrium anticipate if there is a profitable reaction to his defection which will be loss-making for him. If there is he refrains from the defection. Thus, in a separating equilibrium, a potential defector with a pooling contract will anticipate that a separating contract reaction which attracts the high quality types will induce losses on his pooling contract, and will refrain from offering the pooling contract. The original separating equilibrium is then a reactive equilibrium. Instead of modifying the equilibrium, Riley [1985] alters the parameters of the model to reduce rewards to defectors from an equilibrium. If the proportional rate at which the marginal cost of signalling declines with quality is sufficiently large, the profitability of NE-breaking competition is eliminated.

These methods of obtaining existence are not very satisfactory. Model parameter alteration suggests testable conditions under which screening is feasible but seems a very specific way of dealing with a problem common to a variety of markets. The need to change equilibrium concepts according to the existence requirements of particular markets suggests a lack of understanding, rather than explanation, of behaviour. In addition, the active strategic behaviour implied by these notions seems inconsistent with the perfectly competitive structure generally assumed[13] It

---

[10] Insurance companies and banks offer contracts.

[11] In a proposed separating equilibrium with a low proportion of high risk types, some bank will find it profitable to deviate with a contract which attracts all types - a pooling contract. But in a proposed pooling contract, some bank could offer a contract which attracts the low risk, inducing losses on the pooling contract.

[12] Wilson's [1977] anticipatory equilibrium was the first such equilibrium notion.

[13] The two are reconciled by implicit appeal to the Bertrand game of price setting firms with constant re-

implies that markets with asymmetric information must be treated as non-Walrasian.

Two approaches to resolving the nonexistence problem which address these issues or criticisms are those of Gale [1987] and Hellwig [1987]. Hellwig uses a sequential formulation. He suggests that the game theoretic specification (the order of moves, dynamic structure) may determine the equilibrium outcome expected. The nonexistence in the usual two-stage formulation is less acute in a three-stage game. In screening models Hellwig found a pooling sequential equilibrium to be the most plausible, in contrast to Cho and Kreps', op. cit., separating equilibrium in the signalling game. Hellwig's result provides support for the credit rationing equilibrium in a dynamic situation where banks make contract offers and may eventually (in stage 3) reject borrowers' applications. If, on the other hand, potential borrowers announce their preferred contract variable to banks, Cho and Kreps' analysis suggests that sorting contracts are more likely[14].

Gale's analysis is intended to resolve the existence problem within the Walrasian paradigm of parametric and complete prices. He formulates a dynamic matching game in which borrowers and lenders choose a contract by choosing a market to enter[15] and are matched at random. The lender then quotes a price which may be accepted or rejected. The separating sequential equilibrium is not subject to defection because a lender cannot hope to change the set of borrowers he faces by altering price. While this *ex post*-of-contract-choice pricing avoids the auctioneer it does so by losing the notion of competitive pricing in a manner not dissimilar to the auctioneer itself. It could perhaps be used to suggest that in markets with asymmetric information equilibrium depends on an *ex ante* matching of agent and principal that in some sense does away with price competition. It may be, as Gale notes, that a more insightful view of price formation in markets with asymmetric information would be obtained by use of an oligopolistic model.

The models discussed in the following theoretically precede those of Gale and Hellwig. In the main they challenge the conclusions of Jaffee and Russell [1976] and Stiglitz and Weiss

---

turns to scale.

[14] Note that this could mean that in a single financial market there may be a variety of contract equilibria - large borrowers do approach banks (signalling) whereas small borrowers generally choose among available offers. This provides another possible explanation of observed differential treatment.

[15] There is one contract per market and a market for each contract with a distribution of risk types in each market.

[1981] that the competitive credit market equilibrium is one with rationing. They argue that banks have an incentive to offer sorting contracts, eliminating the need for rationing[16]. We use the categories 'rationing by loan size' and 'rationing by exclusion' to classify both rationing and contract models in order to explain how rationing may be avoided by appropriate contracts.

The possibility of rationing by exclusion

Analyses of the possibilities of rationing by exclusion, or its avoidance, start out from the S-W [1981] demonstration that banks face a quality of loan demand which may be affected by the interest rate they charge. Higher rates of interest may attract borrowers with riskier projects who are more willing to pay a higher rate of interest because they know their probability of repayment to be low (the adverse selection effect). Higher rates may also encourage borrowers to undertake riskier, but higher return, projects because this choice maximizes their net expected payoff (the moral hazard effect).

In the S-W analysis banks are assumed to face borrowers whose projects have the same mean return but differ by a mean-preserving spread[17], banks are unable to identify individual project risk. When raising the loan rate changes the riskiness of the pool of potential borrowers through the adverse selection and moral hazard effects, a bank's expected return is not a monotonic increasing function of the loan rate. There may therefore be a bank-optimal rate at which the banks' expected return is maximized. If this is less than the market-clearing rate the loan rate is not raised to eliminate excess demand.

Bester [1985] took issue with the pooling equilibrium found by S-W. He argued that banks could offer several contracts, each specifying different combinations of collateral and the rate of interest, in order to induce self-selection among borrowers. In his [1987] paper Bester considered collateral both as a self-selection and incentive device[18]. If the single crossing property[19] is

---

[16] To the extent that contingent contracts complete credit markets - each borrower/loan type is thought of as a different commodity - no credit rationing would occur.

[17] Riskiness is ranked by second order stochastic dominance (SOSD): a distribution F is said to be less risky than a distribution G in the SOSD sense if the probability of low values of the random variable is greater under G. If G can be generated from F by redistributing probability weight from the centre to the tails of the distribution leaving the mean unchanged, G differs from F by a mean-preserving spread. Distributions with similar mean can then be compared by variability alone (Rothschild and Stiglitz [1970]).

[18] S-W had argued that collateral could have adverse selection effects if wealthier borrowers with more

satisfied and the LR are not constrained in the amount of collateral they can provide, collateral permits banks to sort and provides incentives to undertake less risky projects. However, these advantages are bought at the cost of inefficient risk-sharing, if borrowers are risk averse and banks risk neutral. Further, projects may be riskier than with asymmetric information because less risky choice requires excessive risk-sharing inefficiencies. Besanko and Thakor [1987a] find that when wealth levels are insufficient to permit perfect sorting, it is the low risk (LR) that are rationed. Intuitively, since the LR cannot prove they are LR by collateral, and the high risk (HR) will pay a higher rate of interest, it is the latter who are granted loans when banks cannot sort.

Collateral screening contracts may be less feasible in LDCs where collaterizable wealth is limited - encouragement of the accumulation of such wealth was one of McKinnon-Shaw's arguments for higher deposit rates. However, in some societies, the incentive role of collateral suggests that the market value of collateral is not the only relevant factor. Empirical intuition is lent to Bester's assumption guaranteeing that initial wealth does not constrain collateral, viz. $U(0) = -\infty$, by the observation that the traditional (non-pecuniary) value placed on their plots by farmers in some African countries is so high that some banks consider its acceptance as security as a near-certain guarantee of repayment. These are market features that will vary not only from country to country but temporally if, for example, such farmers' tastes change after long contact with other mores.

Assuming debt contracts with devices similar to those in Bester's [1987] paper, Besanko and Thakor, op.cit., obtain similar results. However, they alter the definition of risk: safer projects first order stochastically dominate (FOSD)[20] the less safe. This change in risk definition is important because, as Besanko and Thakor point out, with FOSD increases in the rate of interest

---

collateral are less risk-averse. Wette [1983] extended the argument to risk neutral borrowers. Neither considered incentive compatible contract offers.

[19] i.e. The LR are always willing to accept a greater increase in collateral, C, for a given reduction in the rate of interest, r, than the HR, because they have a lower probability of losing that collateral.

[20] A distribution F is less risky than a distribution G in terms of FOSD if the cumulative probability of the random variable with distribution F exceeds that of the variable distributed G. The expected value of the F-distributed variable is therefore always larger. In this case less risky projects also have higher expected value, while in the S-W SOSD with a mean-preserving spread, projects with similar expected value are ranked by variability (risk) only.

cause HR borrowers to exit first. Higher rates of interest increase the return required by potential investors and since riskier projects have lower expected returns under FOSD [21] they are the first to leave[22].

The first order stochastic dominance risk definition is also used by de Meza and Webb (M-W) [1987] to find a pooling equilibrium in a competitive economy. There is no credit rationing in this pooling contract because, given the risk definition, banks' expected profits are monotone increasing in the interest rate: when the loan rate is increased the HR exit first. Their model's main interest lies in its consideration of tax policy and it will be discussed at greater length in Chapter 8 where we focus on policy issues.

In their [1986] paper S-W took up the sorting issues raised by the authors discussed above. They argue that the increased contract dimensions provided to banks in these models do not necessarily eliminate credit rationing, given sufficient borrower and technique heterogeneity. Heterogeneity implies that the single crossing property may not exist. For example, while decreasing absolute risk aversion (DARA) implies that the poor's MRS between collateral and the rate of interest exceeds that of the rich, DARA also implies that the rich may undertake riskier projects. Thus the relative slopes of poor/rich indifference curves are ambiguous. Further, as collateral increases all individuals undertake less risky projects. As a result, their MRS vary, and indifference curves between collateral and the loan rate are not quasi-concave and may cross more than once. There may be various degrees of pooling and separating equilibria, both with and without rationing; including a (partially) separating equilibrium [23] where borrowers are rationed at every contract [24].

---

[21] With second order stochastic dominance low risk borrowers, who are those most likely to have to bear the increased cost, are the first to exit when the loan rate increases.

[22] B-T also consider the case of a monopoly bank. Its optimal contract is pooling when the social surplus of the HR is sufficiently large, because this contract enables it to extract maximum surplus. The contract achieves the first best outcome except for a transfer from the monopolist to the LR/high return who is paying a lower rate than he would under symmetric information when the monopoly bank is able to extract all his surplus.

[23] This occurs when there is excess demand at a separating equilibrium but an increase in the deposit rate would require loan rate changes such that contracts on offer no longer sort.

[24] S-W [1987] used this partially separating equilibrium to reply to Riley's [1987] criticism that rationing is unimportant in the macroeconomy since aggregation over all banks leaves only one pool rationed.

We note that all of these authors found either a Nash equilibrium or pointed out that, while the NE might not exist, their model would have a reactive equilibrium.

In addition to sorting devices, banks have direct ways of obtaining information and controlling behaviour: accounting procedures and (monitored) contract specification of loan use are examples. Less obviously, banks can acquire information about a borrower (or his ongoing activities) in the course of a repeated relationship. S-W [1983a] suggests that the latter is itself a source of credit rationing - the fact that the borrower will need the bank again provides the bank with an incentive-giving device in the form of the denial threat. It is always possible that there are some individuals against whom this threat is ineffective - the cost of repayment may outweigh the costs of exclusion, or the individual may have no further need to borrow. Knowledge of individual circumstances may lessen banks' exposure to default risk on some subset of the market. Section 7.4's model and Chapter 8 argue that this may result in asymmetric information being more important for some sectors. Section 7.3 suggests that this may help explain credit allocation in the Caribbean.

All the models considered so far assume that loans are made by financial intermediaries and are governed by debt contracts which take a particular form: fixed repayment if the project is successful, and lender recuperation of realized gains (plus collateral, where applicable) in the event of default, although credit rationing is a consequence of the divergence in the expected returns to intermediary and borrower determined by the contract. Given economic theory's usual prior that individual maximization in a freely operating market will (abstracting from all the required caveats) evolve efficient mechanisms, the question arises of why the market constrains itself to operate with these particular schemes. Williamson's [1987] analysis yields both the standard debt contract and financial intermediation as the endogenous market response to *ex post* [25] asymmetric information between borrower and lender. The debt contract ensures truth telling; intermediary lenders reduce monitoring costs and pool risks. But the probability of costly monitoring increases with repayment size so there may be a bank-optimal loan rate where rationing

---

[25] That is, information about the realized project outcome. His model can be compared to and merges the analyses of Gale and Hellwig [1985] and Diamond [1984]. These were both discussed in Chapter 2.

occurs.

The possibility of rationing by loan size

The analyses discussed above all focus on rationing of the type where individual borrowers (or groups of borrowers) are unable to obtain loans, and where the differences between borrowers arise from differences in the risk characteristics of their project or choice of projects. Jaffee and Russell [1976] (J-R) defined credit rationing to be the supply of a loan size smaller than that demanded at the quoted rate of interest, and took the default risks faced by banks to be risks dependent on the honesty of the individual, the dishonest being those with lower default costs. As loan size and loan rate increase there is adverse selection of defaulters as the proportion of those who find default worthwhile increases. Therefore J-R suggest that competition leads banks to market a pooling contract at the zero profit loan rate, because it attracts the honest.

Milde and Riley's (M-R) later analysis [1988] derived a single pooling contract with no default as a reactive equilibrium in J-R's model, if the proportion of potential defectors is sufficiently high. The contract entails rationing by loan size, as proposed by J-R, because default is avoided by keeping loan size at the level where no borrowers wish to default. Screening, though feasible, is not desirable because it would involve potential defaulters choosing larger loans on which they would then default.

M-R's own model(s) consider informational asymmetry resulting from quality differences in a neoclassical production function, unlike the models considered above where technology was indivisible. Its general form is

$$\tilde{x} = Q(\theta, L, \tilde{u})$$

where $\tilde{x}$ is gross stochastic project return, increasing in L, L is loan size, θ is a privately known quality parameter which varies across projects, and $\tilde{u}$ is a random variable unknown *ex ante* to borrower and bank. Loan contracts specify loan size and and rate. M-R consider three models differing in how the random factor enters the production function. Since we have adopted their neoclassical specification, Section 7.4 compares their model results with our own. Here we note only that in their models 1 and 2, with the FOSD definition of risk[26], the marginal borrower who

---

[26] M-R's production function approach has separable risk and quality, but return is increasing in quality,

departs is the low return, HR whose expected payoff decreases, but in Model 3 with SOSD and similar means, the LR exit, so that the average quality of investment declines.

Besanko and Thakor's (B-T) [1987b] generalize the models above by considering a continuum of types with projects whose returns are a continuous positive function of investment size and success probability. Credit contracts specify loan size, collateral, loan rate and credit granting probability. In the screening contract reactive equilibrium B-T find that the lower risk borrowers[27] signal their success probabilities by accepting a larger loan with a higher rate of interest than they would under symmetric information. This is a case of overinvestment similar to that in Milde and Riley's Model 1 and contrasts with the underinvestment found in our model. The difference, as we will discuss below, stems from the definition of risk used, as does the overinvestment and underinvestment pooling equilibria in de Meza and Webb, op.cit., and Stiglitz-Weiss, respectively. Collateral in this model does not sort but serves to provide the banks with full insurance on the loans of the higher risk whatever the distribution of information.

Credit rationing by loan size is also predicted by Gale and Hellwig (G-H)[28] [1985] as a result of the standard debt contract which responds to *ex post* informational asymmetries between creditor and debtor, and the technology. Given asymmetric information about *ex post* project return and the need for financing, the optimal mechanism is the standard debt contract: its provision of bankruptcy (observation by the lender) and a fixed repayment is required for truth-telling; but bankruptcy is costly and diminishing returns to investment increases its probability as loan size increases. Credit rationing limits this probability. This suggests that, with *ex post* information asymmetries, credit rationing reduces the expected cost of the incentive compatibility requirements of the debt contract, rather than being the result of a failure to write appropriate contracts, as is assumed with *ex ante* asymmetries (compare Williamson's, op.cit, rationing result).

---

and all projects are subject to the same stochastic component.

[27] High risk (HR) and low risk (LR) are borrowers with success probabilities below and above, respectively, a cutoff probability at which full information equilibrium loan size equals wealth discounted at the riskless rate.

[28] Their model, concerned with the more general issue of the firms's financing in incomplete markets, is discussed in Chapter 2.

Theoretical predictions about credit market outcomes

Having examined a range of theoretical models we may ask whether they predict the empirical outcome we should expect following liberalization. We suggest that the modelling of risk, separating conditions and strategic behaviour require that theoretical insights be combined with empirical knowledge to arrive at economy-appropriate analysis.

Comparison of B-T [1987a,b], M-W [1987] and M-R [1988] with S-W makes it clear that the risk concept used has a crucial effect on the outcome of banks' credit policy, whether or not sorting is possible. This arises from the effect an increase in the loan rate has on borrower profit - which determines which risk category supplies the marginal borrower and hence the direction of the adverse selection effect. If expected returns are similar across projects, borrowers with less variable returns have lower expected payoffs and will include the marginal borrowers who exit when the loan rate increases[29]. If borrowers with higher expected returns have higher success probabilities it is the HR[30] who have lower expected payoff and contribute the marginal borrowers who exit with an increase in the loan rate. While the authors recognize this difference in formulation, they do not discuss whether one, or which, assumption should be considered more plausible. It seems reasonable to assume that real banks are confronted with a whole range of potential borrowers, some of whom have equal expected returns, and some of whom do not. The question is then whether banks are able to classify the groups[31]. Ignoring auditing (monitoring) possibilities, banks would have to sort between pools with different expected returns, as well as within each pool.

The models demonstrate that it is in principle possible to avoid credit rationing which arises from *ex ante* informational asymmetry if the bank has sufficient degrees of freedom relative to the decisions or characteristics being controlled or screened for (Hart [1986]); and if the contracts are noncooperatively enforceable on the market (Hellwig, op.cit.).

---

[29] Expected payoff functions are convex in project return and the distribution of the HR is a mean-preserving spread (MPS) of the distribution of the LR. The expected value of a convex function is increased by a MPS and therefore LR expected payoff is less than that of HR.

[30] Those whose projects are first order stochastically dominated.

[31] S-W assumed that banks could identify pools by their expected returns.

To take the first point: all the models consider the default risk attached to any borrower/project as varying in only one dimension. This limits the reliability of their predictive implications. Paraphrasing Engers [1987], when differences are multidimensional, potential borrowers can no longer be simply ordered in terms of their marginal cost of signalling (marginal rate of substitution between price and the signal). One borrower may have a lower relative marginal cost for some signals and a higher relative marginal cost for others. We can imagine, for example, borrowers differing in both management ability and honesty, with project techniques of varying riskiness. These differences could be especially important in the lagging economies of McKinnon-Shaw, where frustrated but profitable investment opportunities are viewed as originating mainly in the informal household sector and it is easier to confound individual and activity characteristics: default risks arise from the capability of the household-firm, the honesty of the borrower and the project technique adopted. There is also a moral hazard problem because a bank would find it difficult to control the purpose for which a loan is used: consumption or investment. One can always constrain a particular model so as to ensure that a bank has sufficient instruments, or add enough variety to ensure that it does not - that adverse selection remains a source of rationing; for the purposes of policy analysis, empirically based judgement seems the best basis for the choice.

There is a more fundamental aspect. For sorting to be feasible it is assumed that banks know the distribution function of the risks on the market. And the production functions of the borrowers. While this may be a reasonable assumption in industrialized economies where banks have developed an acquaintanceship with the market,[32] it seems less plausible in LDCs where urban-based banks, with staff trained in loan assessment of traditional commercial activities, may be receiving loan applications from rural applicants and small-scale urban borrowers operating outside of the 'formal' sector. McKinnon's 'suggestion' that moneylenders be recruited as loan officers in commercial banks recognized that this could pose a problem. McKinnon-Shaw's insight that information-process failure is an important contributor to underdevelopment, suggests

---

[32] Except perhaps in the case of innovative industries for which venture capital financing has been developed.

that more stress should be placed on such issues in LDCs.

Consider the second point: the Hellwig [1986] and Gale [1987b] examinations of rationing further suggest that whether pooling or separating is predicted depends on the dynamic structure and order of moves of the game played. Their analyses also raise questions about behaviour on these markets that are not easily answered. Hellwig explains his results by a tension between the vulnerability of pooling contracts and Bertrand competition. A competitive bank offering a pooling contract runs the risk of losses if the high quality types leave, but Bertrand competition leads towards return and cost equalisation at the margin - not separation. Which of these 'forces' dominates depends on the proportion of high quality types and the order of moves, and determines whether or not there is credit rationing. There is, it seems, a conflict between competitive pricing and the strategic behaviour required to plan and implement contracts that sort. This has parallels with Gale's work where contracts had to be taken as given, and the device of matching with *ex post* pricing used to achieve a Walrasian equilibrium.

These models indicate that the existence of asymmetric information calls forth a variety of private mechanisms which can permit an incentive compatible transfer of the information. In this sense they support the McKinnon-Shaw intuition on the informational role of banks. But, as we will see in Chapter 8, implementation of these mechanisms is costly. These analyses also suggest that an observed market outcome can be an amalgam of responses whose sources and effects are not self-evident; we will see in our model that screening by loan size can produce loan allocations that are probably observationally indistinguishable from rationing.

In the model of Section 7.4 we attempt to address one of the points raised here by adopting the SOSD risk definition but not mean-preserving spread; expected return and risk can then vary in a non-monotonic fashion since return also depends on a management quality variable.

In Section 7.3 we consider observations and comments on credit markets in our four Caribbean economies, to see how these relate to the priors suggested by this literature.

7.3. Theory and Practice in LDCs

Credit allocation in LDCs is usually thought to be suboptimal: loan returns are not equalised at the margin across sectors because some sectors or activities receive smaller loans than they should, or receive no loans at all. Output would be higher if credit flows were directed to more productive sectors. Attributing this outcome to regulatory distortion, current thinking favours minimum regulation. It has its origins in the McKinnon-Shaw analysis of financially repressed markets. But policymakers in LDCs have until recently taken the opposing view that society's interests and those of the private creditor diverge and that regulatory suasion may be required to reconcile the two.

Recognition of private information raises theoretical doubts about the probability of an efficient outcome in a liberalized credit market. Agents in a regulation-free market will continue to possess information about themselves and their activities that creditors do not have, and will act privately in their own interests. Profit maximizing creditors must condition their decisions on this knowledge. The present section makes the following points. First, the observations from which credit rationing is usually inferred, as illustrated by outcomes in the Caribbean markets, are consistent with rationing of the Stiglitz-Weiss varieties, may be consistent with McKinnon-Shaw regulation-induced rationing, and even with screening contracts. Care must therefore be taken in inferring arrangements from simple observations. Second, the issue is far from trivial, given the importance of debt financing in LDCs. We take the last point first.

External Financing of Enterprise

In industrialized countries distortions in the credit allocations made by financial intermediaries may be of second order importance outside of particular sectors. Firms have feasible access to alternative sources of external financing[33]: public share issues, bond issues, venture capital. Thus, if rationed out of the bank market they may be able to substitute other funding (credit or equity) sources. McKinnon-Shaw however drew our attention to the pivotal role financial inter-

---

[33] We abstract from the asymmetries of information which also affect these alternatives - as suggested by, for example, Leland and Pyle [1977].

mediaries' credit[34] allocation plays in aggregate real economic activity in LDCs. As virtually the only source of external finance, financial intermediary loans affect production possibilities, employment, investment demand, technological choice etc.

In the Caribbean, apart from some twenty years of stock exchange operations in Trinidad, and a nascent exchange in Barbados, alternatives for external financing consist of commercial banks, finance companies, trust companies, insurance companies, trade credit, pension funds and development banks. Among the countries considered, only Belize has an unorganized money market worth mentioning. Therefore, with the possible exception of small-scale rural enterprise in Belize, it can be presumed that there is no curb market to which rationed firms have access. We consider the alternatives to private bank debt available in the Caribbean.

Development Banks

The development banks were officially established to provide long-term credit to producers, the credit allocation to be determined by project evaluation[35]. Since their objective function is presumably the social return to the project (i.e. both borrower's and bank's), and they are mandated to investigate, the informational asymmetries which affect private decisions should be less determinant in development bank financing. However, two factors suggest that they do not greatly expand the options available to potential borrowers. Firstly, the non-profit maximand itself appears to create bureaucratic procedures which raise the transactions cost of their credit[36]. Secondly, the development banks depend on state subventions, external borrowing and bond

---

[34] Usually taken to be commercial bank credit for the simple reason that commercial banks dominate the financial sectors of most LDCs.

[35] There are two ways in which official project evaluation and monitoring appear to be able to overcome some of the informational problems that necessitate sorting by equilibrium credit rationing or contract screening. Primo, suppose entrepreneurs are unwilling to risk losing their competitive edge by entrusting their productive ideas to a profit-maximizing institution who may exploit them. The institution may not be able to obtain sufficient information to decide on a loan. Official indifference to profit avoids this moral hazard problem. (This argument of course ignores unprincipled behaviour on the part of functionaries). Secundo, part of the rationale of development banks has been the supposition that inexperienced entrepreneurs are unable to undertake the feasibility studies necessary to 'certify' their project. A private commercial bank may not have the incentive to undertake detailed and costly evaluation. If the project is found to be non-feasible, it will have sunk the evaluation costs unprofitably. However, a successful evaluation requires the active assistance of the entrepreneur who is thus furnished with the information required to certify the profitability of his project. The entrepreneur then has an incentive to use this information to request credit from another bank. The latter can charge him a lower rate of interest since it does not have to cover the costs of project evaluation.

[36] Although their lower loan rates may provide some compensation.

finance for loanable funds. Their funding capacity is therefore very limited in comparison with private financial institutions. For example, in Belize in 1981 the Development Bank held only 14.4% of financial institutions' assets and these were in part allocated to official institutions[37]. Similarly, in Barbados in 1982, the Development Bank contributed only 8.4% of total long-term credit, which was in any case only 27.2% of total business borrowing[38].

Stock Markets

The equity market is usually assumed to be the principal source of non-debt finance. However, Greenwald, Stiglitz and Weiss [1984] argue that it is itself plagued by informational problems, which may reduce its ability to substitute for denied bank credit. If stronger firms are more willing to assume the increased risk of greater debt, equity issue may signal poor firm quality, dissuading potential issuers. In this view, asymmetric information reduces the availability of external capital, whatever its source. On the other hand, Cho [1986] argues that an equity market may permit financial liberalization to achieve an efficient capital allocation. He suggests that the adverse selection and moral hazard phenomena which produce bank credit rationing do not affect equity finance since the equity investor has the same expected return as does the project. Apart from the Greenwald-Stiglitz-Weiss adverse selection argument above, this ignores the potential moral hazard problem between firm management and equity investor. Leland and Pyle [1977] also suggest that adverse selection may deter high quality projects. These difficulties of credible information transfer may account for the relatively non-active nature of even the established stock exchanges in the Caribbean. The stock exchange in Trinidad serves mainly as a secondary market, but even its secondary trading is insufficient to provide adequate liquidity prospects for investors. In the ten years to 1978, equity declined steadily from 36 to 19 per cent of local corporate liabilities. (See Bourne [1982], Table 11). Bourne attributes this to reluctance to divest family or foreign control, to substantial issuing costs and to marketability problems on the demand side. Most activity on the stock market during these years was in fact generated by equity sales by foreign firms responding to government promotion of local ownership (see

[37] See Table 4.1 and Luben [1983], pages 23-26.
[38] Worrell and Prescod [1983], page 15

Bourne [1982] and Ramkissoon [1981]).

Insurance companies and finance companies

Only the private financial intermediaries remain to be considered. Table 7.1 shows the institutional distribution of financial firms' assets for the four countries being considered[39]. After commercial banks, insurance companies, finance companies and trust companies are the principal financial institutions. In the UK and USA, insurance companies are a major source of finance for the corporate sector through their holdings of either equity or bonds. Of course, the types of small-scale risky enterprise considered important in the LDC context are not likely to furnish a suitable investment for insurance companies[40], but it is clear that insurance firms provide little capital, even when we consider the total business sector. In Trinidad and Tobago, Table 7.1 indicates that they sourced only 5.4% of the local corporate sector's funds. In Barbados, corporate shares averaged only 6.1% of the local assets portfolio[41] of all insurance companies between 1979 and 1983[42]. Trust companies and finance companies (the latter in the Bahamas) play a role similar to building societies and savings and loan associations in the UK and USA. They specialise in long-term residential mortgages and construction/real estate investment. In fact, in several cases the trust companies are subsidiaries of commercial banks established to cater to the long-term requirements of the household sector. In the Bahamas, 73.5% (in 1978) of finance companies' assets were held in the form of mortgages[43]. Investment in shares was negligible. In Barbados published trust company balance sheet information does not include an equity holdings

[39] With the exception of Trinidad and Tobago, which has flow-of-funds data, these shares would not be entirely accurate since they represent proportions of summed financial assets. In Trinidad's case, the institutional structure of local liabilities is used because their flow-of-funds financial assets information includes the central bank and government, and does not provide the shares of non-bank financial institutions.

[40] In these micro economies which are, moreover, geographically close and economically linked, the diversification required by insurance companies is likely to be a more crucial issue for domestic or regional companies.

[41] Source: Table D22 in Central Bank of Barbados [1988]

[42] No information is available on insurance companies' portfolios in Belize. Statistical information for insurance companies in the Bahamas is too aggregated to permit a similar comparison. However, real estate and mortgages appear to constitute their major form of long-term investment, averaging 46.1% in 1973-1977 (Table 4.20 in Ramsaran [1984]). Ramsaran's investigations, op.cit., p.178, indicate that most of these mortgages were for residential rather than business purposes.

[43] Source: Table 4.2 in Ramsaran, op.cit.. While published data for the Bahamas does not provide a distribution of mortgages by sector, viz. residential, commercial or industrial, Ramsaran again (op.cit., p.225) indicates that most of the listed mortgages are in fact residential.

category. But these would have been registered in the 'other assets' category which was less than two percent of total assets in 1983. Only 11.9% of total trust company loans were extended for industrial and commercial purposes in that year[44].

Commercial banks' equity investment

Commercial banks do not appear to invest in private securities. These averaged less than two percent of commercial bank assets in the Bahamas between 1974 and 1978[45], and less than one percent in Barbados in the five years to 1983[46]. Belize's published bank balance sheet data lists no private securities. Trinidad's banks held an average of only five percent of assets in *both* 'government' and 'other' securities between 1976 and 1980; given reserve requirements, it is not unreasonable to suppose that these consisted mainly of treasury bills. Commercial banks also discount commercial or trade bills for business firms, providing an indirect source of finance. However, in Trinidad, the only country for which data is available, such bills only averaged 1.3%[47] of total assets between 1976 and 1980.

Other sources

Other possible sources of external financing are pension funds and credit unions. With the exception of Trinidad and Tobago, no information on pension funds is available for any of the countries. In Trinidad pension funds held less that five percent of the corporate sector's liabilities, and credit unions held a similar proportion of the households and unincorporated business sector's liabilities (see Table 7.1 and Bourne [1982]). In Belize, where credit unions may conceivably provide the finance required for small-scale business, membership is a prerequisite. Members are required to share a common characteristic: geographical location, economic or occupational[48]. They are therefore not a general source of finance although they are an important source of credit for rural areas, and the bond between members may serve to mediate informational asymmetries in areas where modern urban financial institutions are at a particular

---

44 Source: Table D1, Central Bank of Barbados, op.cit.
45 Source: Table 2.9 in Ramsaran, op.cit.
46 Source: Table B1 in Central Bank of Barbados, op.cit.
47 Source: Data on Trinidad's bank portfolios is from Table 2 in Ramkissoon, op.cit.
48 Luben, op.cit., page 18

disadvantage.

This description of the financial assets in these economies indicates that business finance is largely limited to that available from intermediaries, and that their finance generally takes the form of debt. It is therefore subject to agency constraints of the type analysed by the models discussed in Section 7.2. If, in response to these agency problems, intermediaries ration credit (the interest rate is the screening device), or issue debt with screening variables set such that some potential borrowers are unable to obtain credit, or receive a loan other than that they would have received under common information, borrowers' realized production, investment demand etc. may differ from their notional quantities. Blinder [1987] has made a first attempt at formalizing the effects of equilibrium credit rationing on the macro front. In his model(s) firms' only source of working capital is bank credit which is rationed because of informational asymmetries. The functioning of the macro economy depends on whether or not the credit constraint is binding. If demanded credit is unavailable to producing firms, 'there may be a "failure of effective supply"' where firms are unable to produce their notional supply. As economic activity expands, default risk declines so that banks are able to reduce excess reserves and increase credit. In the working capital model with credit rationed, monetary policy (which raises bank reserves)[49] has an expansionary effect on output because customer credit availability is increased.

The idea of effective supply failure due to credit unavailability has often been assumed by Caribbean analysts: viewing commercial bank credit as the 'main source of working capital', their models have attempted to

> "provide explanations which correspond to observable behaviour, by focussing on credit-expenditure rather than money-expenditure relationships"[50].

This scenario differs markedly from that of McKinnon-Shaw who attributed credit rationing to repression. They recommended the minimization of regulation, particularly the removal of limits on interest rates, to correct the effects of rationing on output. In a Blinder-type regime such

---

[49] Note that the more usual channels of monetary policy, via interest rates, for example, are assumed away; money, with the crucial exception of bank reserves, has no 'essential' role.

[50] Worrell [1985]

deregulation would have no effect.

Credit Rationing after Liberalization?

The credit rationing described by McKinnon-Shaw and by Caribbean analysts appears closest to the Stiglitz-Weiss rationing of observationally distinct borrowers[51]. In the latter banks are able to distinguish groups of borrowers, where members within each group have a common observable attribute. But borrowers in the same group (where the group may be taken as an industry, sector, class etc.) may differ in risk and productivity, these being determined by both their observed and unobservable attributes. A bank's gross return from lending to each group is concave in the loan rate, just as it is concave in the loan rate charged to a market of unidentifiable individuals. Rationing may therefore occur for similar reasons, but whole groups may be excluded. The point here is that the expected social return on loans to those excluded may be greater than the return to those receiving loans, since the bank is concerned only with the return in the event of success.

McKinnon's and Shaw's Views

A major motivation for McKinnon and Shaw's recommended liberalization of the financial sector in LDCs was the generation of loanable funds for the financing of investment projects with high real rates of return. They characterized repressed economies as rich in high-return investment opportunities (Shaw, p.81)[52] whose realization is frustrated by the rationing of potential borrowers. Rationing is attributed to regulation on both the supply and demand sides of the market. Regulated low deposit rates (in conjunction with high variable inflation resulting from high nominal money growth rates) reduce the real supply of loanable funds (savings) to the banking system. Usury ceilings on loan rates prevent banks from raising rates to the levels necessary to compensate them for the information acquisition costs and risks accompanying small-scale lending to non-traditional sectors (McKinnon, p.73). Further, the low loan rates permitted to

---

[51] Carter [1988] also used this analysis to argue that unrestricted markets may ration credit to small risky agricultural borrowers.

[52] Page numbers refer to Shaw's [1973] and McKinnon's [1973] books.

banks reduce the rates they can offer to depositors (McKinnon, p.69). At these regulated interest rates there is an excess demand for credit, the available loanable funds supply being allocated not by price, but by the decisions of government officials and the banks' constrained preferences for established low-risk borrowers (Shaw, p.86; McKinnon, p.68). By designating priority sectors/entrepreneurs who receive exclusive licenses (for imports, for example) the authorities effectively reduce the risk attached to those loans, and banks themselves restrict their lending to "completely safe borrowers whose reputation is known, or whose collateral is relatively riskless" (McKinnon, p.73), including government itself. The result of this interaction between regulations and banks' optimizing behaviour is that small-scale and/or rural borrowers have access only to the unorganized money market: moneylenders (the curb market) and the village store, whose unregulated interest rates reflect the excess demand for credit and the risks attached to evading regulation (Shaw, p.89), as well as the monopoly power acquired by operating with specialized market information in dispersed localities (McKinnon, p.72[53]) As a result, moneylenders' rates are above even the high potential rates of return to investment (Shaw, p.122; McKinnon, p.6) and discrete expenditures required for innovative investments are blocked by the constraint of self-finance (McKinnon, p.12).

Caribbean views and policy responses

While Caribbean economists and policymakers also identify specific groups as being excluded from the market, they see such rationing as symptomatic of the limited nature of banks' intermediary role. It is argued that banks tend to refuse loans to small-scale and/or non-traditional activities and investors. This tendency is attributed to a high degree of risk aversion on the part of banks, the inability of excluded borrowers to provide collateral[54] and the banks' reliance on their special knowledge of particular sectors. The asset portfolios of the financial institutions reflect both the main areas of economic activity and traditional loan patterns. There is a historically-based dependence on imports throughout the countries of the region. As a result

---

[53] Although Shaw (p.89) argues that free entry to curb markets reduces lenders' monopoly power.

[54] A theoretical rationale for the collateral presumption is provided by the models in which banks sort among their credit customers through collateral requirements.

domestic "big" business is concentrated in distribution. In countries whose productive mainstay has been sugar, traditional sugar industry links are also a determinant of size and creditworthiness.

Manufacturing, agriculture and tourism are seen as the sectors with most growth potential. The distribution sector (retail and wholesale trade) is viewed as contributing little to growth, especially since its activities are mainly import-related in economies where the balance of payments is the crucial constraint. Thus in all the countries, the observationally distinguishable groups viewed as being rationed are small-scale/innovatory manufacturers, non-traditional agriculture etc. This view is not dissimilar to that of McKinnon, who discussed the rationing of the poor farmer and 'small domestic machine shop' operator. The penchant among banks of the area to lend to the distribution and personal sectors is seen as impeding the realization of new productive structure. As a result, in the 1970s Trinidad and Barbados introduced selective credit controls aimed at restricting the growth of commercial bank loans to the distribution and/or personal sectors. In Barbados attempts were made to reinforce these by guaranteeing, or offering rediscount schemes for, loans to the 'rationed' groups. With the exception of the sugar industry, banks have not taken advantage of the latter. The data in Tables 7.2 and 7.3 reflect the results of such credit restrictions. Barbados has also had in place extensive interest rate controls. However, interest rates in the other three countries were market-determined. Neither the Bahamas nor Belize have attempted to fix credit guidelines.

Selective credit guidelines were introduced in Trinidad and Tobago in 1979 in an attempt to restrict personal sector credit which grew rapidly following the petroleum boom. They have not been fully successful: while individuals' borrowing averaged 44 percent of total loans in the three years before 1979, at 35 percent in 1982 it was still some way above the 25 percent target. Bourne ([1984], p.43) attributes this large share of personal lending to banks acting "to ration credit to particular sectors or categories of borrowers". He explains this by S-W adverse selection and by the constraints placed on appropriate pricing by oligopolistically-determined rates. In Belize personal sector credit is only about 10% of the total, but the distribution sector and sugar-based agriculture are the favoured borrowers. About half the loans to agriculture are in fact to

sugar-growing, and even the credit to manufacturing shown in Tables 4.2 and 4.3 mainly represents loans for sugar-processing (Luben, p.29). Barnett's ([1983], pp.143-144) description of the banks' loan criteria suggests a possible explanation in terms, for example, of Bester's [1987] and/or Besanko and Thakor's [1987a] models:

> "...the most important requirements are adequate collateral and/or a guarantee and a sound credit record. The banks prefer as collateral mortgages on urban property, insurance policies, fixed deposits and papers of ownership. ... (They) hesitate to accept ... mortgages on small rural land holdings... Consequently, loans to small farmers, with the exception some small sugar cane farmers, are very rare."

In effect, one could consider the banks as offering imperfectly sorting debt contracts with collateral serving as a signal (vs. adverse selection) and incentive device (vs. moral hazard). Banks' limited knowledge of the rural economy makes it difficult for them to judge the value a potential borrower attaches to rural property. Then only borrowers in sectors where the bank has sufficient *a priori* knowledge and where borrowers are not wealth -constrained will receive loans. And, as discussed by Besanko and Thakor, on the assumption that he knows the risk characteristics of the borrower, a co-signer, i.e. a guarantor, serves a role very similar to collateral in avoiding rationing.

A guarantee of a different sort may have influenced the loan portfolio of banks in the Bahamas where, as Tables 7.2 and 7.3 show, real estate and construction receive a quarter of loans. This reflects 100% public guarantees provided to promote housing loans to land-owners (Ramsaran, op.cit., p.96). Ramsaran (p.93) also observes that a "noticeable feature" of the data on commercial bank loans "is the relatively insignificant volume of funds that is made available to the agricultural and manufacturing sectors". Similarly, Worrell ([1985], p.7) argues that:

> "...in the Caribbean customers are rationed by institutional norms, not availability of funds. Those who do not meet conventional credit standards never qualify for loans, no matter how liquid the banks are,"

In Barbados the two sectors *not* subject to rationing are taken to be firms with a distribution sector base and the personal sector. In 1976, before the imposition of credit controls, personal loans had grown to 27% of the banks' loan portfolio. Tables 7.2 and 7.3 show that, despite the selective credit controls imposed in 1977 to limit growth in credit to these two sectors, they continue

to receive the largest shares.

The preference for personal sector loans shown by banks in the richer Caribbean economies could reflect the availability of the collateral signal or incentive device. Personal loans are secured by, for example, mortgages, insurance policies, fixed deposits and the consumer goods for which credit is extended. Any evidence of loan size screening would be purely anecdotal and the S-W [1986] discussion of the effects of wealth and risk aversion on the indifference curves of borrowers (see Section 7.2) clearly suggests why sorting is likely to be highly imperfect. The high degree of regulation of the financial sector in Barbados suggests that country as a possible candidate for M-S type rationing. However, Barbados is the one country where, between them, manufacturing and tourism receive some 30% of credit. Much of this credit is extended to firms which are partially owned by multinational companies or subsidiaries of distribution-based holding companies. The manufacturing category also includes loans to sugar factories. That is, the firms are known to the banks or have the wealth required to provide a collateral signal.

The underlying perception in the Caribbean literature appears to be that financial intermediaries are impeding the structural transformation of the economy (constraining aggregate supply) by concentrating loans in well-established but declining activities when resources should be shifting in line with world prices and in order to diversify the economy's production. If collateral is indeed an important screening variable, it is precisely the non-established sectors who are likely to be wealth-constrained and hence subject to rationing. Shaw (p.86) explains the "privileged place" reserved by banks for "established borrowers, especially trading firms with a long record of stability" by "effective low ceilings on interest rates". The explanation seems inadequate for Belize and Trinidad, as well as the Bahamas, since none of these countries had ceilings on rates. Further, the banks had no shortage of loanable funds.

McKinnon [1988][55] has recently suggested that if banks believe government will rescue them from the consequences of their bad loan decisions they may have less incentive to screen out bad risks. He therefore suggests that an interest rate ceiling (which would act like the bank-

---

[55] See also the discussion in Chapter 8.

optimal rate in the S-W analysis) would correct these moral hazard effects. However, relatively few LDCs appear to offer deposit insurance (we note that Trinidad does) and where there are a large number of multinational banks, as in the Caribbean, their size relative to local economies makes such insurance (explicit and implicit) rather ludicrous.

This suggests that it is not possible to make recommendations that do not discriminate between the conditions in each country. Detailed analysis of the contracts offered by banks and the type of rationing, if any, is required both to determine an appropriate policy stance and to assess its results. For example, McKinnon attributed the substantial growth in real output experienced by Korea after the 1960s to their financial reforms, arguing that the reform relaxed the financial constraint on supply. More recent examinations of the NIC experience suggest that the financial reforms neither liberalized the sector, nor made the assumed contribution to growth. Harris [1985] argues that Korea's 1965 reforms acted to consolidate the role of the state in the financial sector, rather than to reduce its influence. The high rates of interest channelled funds into state-controlled commercial banks, and the organized sector, in which the state could control the allocation of funds. Further, by increasing the range of rates, the reform provided the state with greater scope for differentiation between sectors in charging subsidized rates to selected borrowers (p.23). Harris' view receives support from Park [1988] who argues that Korea's high growth after the 1960s is explained by their export promotion policies - a strategy the state was able to pursue because (Harris, p.11) government obtained the necessary supply of loanable funds to offer subsidized credit in virtually perfectly elastic supply. While Park argues that this degree of government control has produced economic concentration and misallocation of resources, it has also produced the economic growth so many LDCs strive in vain for, and which is the ultimate goal of the liberalization policies. Its applicability in other economies can be questioned in terms of both organizational and political feasibility. The Korean state's success may in part be due to the very degree of their intervention, so that less concern with private incentives was required.

Apart from our warning about the interpretation of empirical outcomes, we take from this section the idea that only particular sectors may need to be screened.

Table 7.1

Institutional Distribution of Financial Assets (percentage)					
Institution	Bahamas (1977)	Barbados (1983)	Belize (1981)	Trinidad (1978) HIE	Trinidad (1978) LCE
Commercial Banks	66.7	69.2	77.5	50.1	30.9
Finance Cos.	17.9	1.6	-	3.7	8.9
Insurance Cos.	13.5	14.1	...	9.0	5.4
Trust Cos.	1.0	7.9	-	sc	2.9
Savings Banks	0.9	-	2.0	-	-
Development Banks	...	4.5	14.4	-	...
Public Fin. Insts.	-	2.8	-	11.4	3.9
Credit Unions	-	...	6.1	4.8	-
Bldg Societies	-	-	-	6.5	-
HIE	-	-	-	-	14.1
ROW	-	-	-	sc	19.5
Other	-	-	-	14.5	14.4

NOTES: - denotes not applicable; ... denotes not available; sc = see country note

BAHAMAS: Bahamian dollar assets. Source: Table 3.16, Ramsaran [1984]

BARBADOS: Proportional distribution of the assets of the listed institutions. Where possible, inter-institutional holdings have been netted out to avoid double-counting. The public financial institution here is the mortgage finance company. Source: Central Bank of Barbados, *Annual Statistical Digest*

BELIZE: Proportional distribution of the assets of the listed institutions. Source: Table 21, Luben [1983]

TRINIDAD & TOBAGO: Financial liabilities of households and unincorporated enterprises (HIE) and the local corporate sector (LCE). As this is flow-of-funds data, the coverage is more comprehensive than that available for the other countries and is not directly comparable.

ROW refers to the foreign corporate sector and the rest of the world.

sc indicates that this category is included in 'other'.

Other for HIE includes the local corporate sector, pension funds, foreign corporate sector, government nonfinancial enterprises, central and local government and trust companies.

Other for LCE includes pension funds, mortgage financial institutions, central and local government.

Source: Tables 9 and 10, Bourne [1982]

Table 7.2

Sectoral Distribution of Commercial Bank Loans (percentage)				
Sector	Bahamas (1978)	Barbados (1983)	Belize (1982)	Trinidad (1982)
Personal	16.4	19.7	8.1	34.5
Distribution	12.3	14.6	31.2	10.1
R. Est. & Con.	24.9	4.0	10.5	16.3
Tourism	14.4	15.4	1.8	0.7
Manufacturing	1.9	15.6	13.7	10.9
Agriculture	0.4	5.0	19.7	1.2
Government	15.1	1.9	10.9	10.3
Other	14.6	23.8	4.1	16.0

NOTES:

General: R. Est. & Con. = Real Estate and Construction; 'Government' includes loans to governments and statutory bodies

BAHAMAS: Loans to residents in both Bahamian and foreign currencies; Sources: Ramsaran [1984] Tables 2.9, 2.14, 2.15 2.14

BARBADOS: Source: Table B6, Central Bank of Barbados [1988], *Annual Statistical Digest*

BELIZE: Source: Luben [1983], Table 14

TRINIDAD & TOBAGO: Source: Bourne [1984], Table 10

Table 7.3

Sectoral Distribution of Commercial Bank Loans Five-Year Average (percentage)				
Sector	Bahamas ('74-'78)	Barbados ('79-'83)	Belize ('78-'82)	Trinidad ('76-'80)
Personal	15.0	22.6	10.2	40.9
Distribution	10.7	14.6	28.4	13.8
R. Est. & Con.	25.6	6.5	11.7	...
Tourism	11.6	12.7	0.9	...
Manufacturing	5.1	15.9	12.2	17.1
Agriculture	0.4	4.2	24.1	2.4
Government	13.6	1.6	3.9	4.4

NOTES:

General: R. Est. & Con. = Real Estate and Construction; 'Government' includes loans to governments and statutory bodies; ... denotes not available

BAHAMAS: Loans to residents in both Bahamian and foreign currencies; Sources: Ramsaran [1984] Tables 2.9, 2.14, 2.15

BARBADOS: Source: Table B6, Central Bank of Barbados [1988], *Annual Statistical Digest*

BELIZE: Source: Luben [1983], Table 14

TRINIDAD & TOBAGO: Source: Ramkissoon [1981], Table 5

7.4. A Model of Bank Screening by Loan Size

In the following model loans supplied by banks are treated as a variable input in firms' production functions. This neoclassical specification allows loan size to be considered as a screening device, and provides a model consistent with a simple general equilibrium framework. It has been adopted from Milde and Riley's [1988] analysis. We find a separating sequential equilibrium in which the low quality firms (indexed by a quality parameter) signal by their choice of a lower loan size (compared to the first best) at the bank zero profit loan rate for their risk class. In contrast, in Milde and Riley (M-R) it is the high quality who signal - by a larger loan and rate of interest in their model 1 and by a lower loan size and rate in models 2 and 3. The differences in these results stem from the following specification changes.

a) We model risk discretely, adopting the normalization that project output is zero in the event of failure. Our zero isoprofit contours are therefore straight lines at the risk-appropriate loan rate factor which is therefore the same under both symmetric and private information[56]. M-R model risk by a continuous random variable, and the bank collects residual output in the event of failure. This simplification allows us to specify the loan rate precisely.

b) We use a SOSD risk definition, similar to Bester's [1987] modelling, with higher quality firms undertaking riskier projects, in line with the assumption that higher return activities are associated with greater risk. M-R's models 1 and 2 adopt a FOSD definition of risk: all projects are subject to the same random factor so the high quality are less risky. Their model 3 has SOSD with a mean preserving spread and low quality are defined as those with the riskier distribution. In general it is the LR who signal their type in order to benefit from a lower rate. In our model this implies that it is the low quality who signal, whereas in M-R the high quality, who are LR in their context, signal.

c) Screening by lower loan size results from the single crossing property of the isoprofit contours. In their model 3 the HR (low quality) is reluctant to accept a smaller loan; in ours the high quality (HR) is unwilling to accept a smaller loan.

---

[56] A similar simplifying assumption is discussed by Clemenz [1986] who attributes it to Bester.

The Bertrand paradigm is presumed to yield a perfectly competitive outcome. This assumption violates the hypothesis of oligopolistic behaviour which, it is argued elsewhere in the thesis, characterises the banking sector in LDC markets. Artificially confining the analysis to Bertrand competition both simplifies it and permits a *ceteris paribus* indication of the effects of asymmetric information on the Pareto efficiency of an otherwise-unsullied credit market.

FIRMS

It is usual in screening models to treat the results as if they apply to the entire economy of borrowers, i.e. all risk classes in the economy are offered screening contracts by the banks. However, in line with the idea, discussed in 7.3, that there are observationally distinguishable groups on the credit market, it is assumed here that banks need to screen only a sector of the potential borrowers. The output y_k of a nontraditional sector or industry is therefore distinguished, and that sector is presumed to be the only one subject to asymmetric information. In other sectors (industries) banks have *a priori* knowledge of each firm.

In the sector or industry producing output y_k, each firm's stochastic output is given by the production function:

$$\bar{y}_k = g(L)V(\theta_f,\mu) \tag{7.1}$$

where L is loan size, $g(L)$ is the production function; $g'(L) > 0$, $g''(L) < 0$. θ_f denotes a firm-specific quality variable, such as firm management. Its value is *ex ante* private knowledge of the firm and it is not a choice variable of the firm. μ is a random variable which takes the values 0, 1 with probability $(1-q_f)$ and q_f, respectively. Its expected value thus depends on the quality of firm management. $V(\theta_f,\mu)$ is thus a function which captures the stochastic effects on output of firm management. The function $V(\theta_f,\mu)$ has the following properties: $V(\theta_f,0) = 0$, all θ_f, by normalization; $V(\theta_f,1) = V(\theta_f)$; $\dfrac{\partial V(\theta_f,1)}{\partial \theta_f} > 0$. In order to permit explicit derivation of factor demand functions and profit functions, the production function, $g(L)$, is specialised to the constant elasticity of output function: L^β, $\beta < 1$.

As is usual, the simplified case with two types is considered. The two types are denoted θ_2

and $\theta_1, \theta_2 > \theta_1$. Their failure probabilities are given by $(1 - q_2) > (1 - q_1)$. Thus

$$E[V(\theta_f, \mu)] = V(\theta_f, 0)(1 - q_f) + V(\theta_f, 1)q_f$$
$$= V(\theta_f)q_f$$

and whether or not the expected value of $V(.)$ for θ_2 exceeds, equals or is less than the expected value for θ_1 depends on the assumptions made concerning the relative values of $V(\theta_2, 1)$ $(V(\theta_1, 1))$ and q_2 (q_1). The equality case occurs when the distribution of $V(\theta_2, \mu)$ is a mean preserving spread of the distribution of $V(\theta_1, \mu)$. Larger θ can be interpreted as indicating better management in the sense that, if successful, any production process undertaken by larger θ will yield more, and have a higher average and marginal loan product. However, θ_2 is also riskier since its failure probability is larger. θ can be thought of as summarizing a Schumpeterian type of entrepreneur-management function. The better quality entrepreneur-manager produces higher returns by using more innovative, and hence riskier, techniques.

Firm f's expected profit is

$$E\phi_f = q_f \left[p_k L^\beta V(\theta_f) - RL \right] \tag{7.2}$$

where ϕ_f is the profit function of the firm with management type θ_f, and p_k is the market price of y_k. $R = (1+r)$ is the interest rate factor. In the event of failure, the firm makes no repayment.

The following standard results are obtained:

1.The loan demand of the applicant for credit is the short run factor demand and is decreasing in the loan rate.

From the first order conditions for profit maximization, loan demand is defined by:

$$\beta p_k V(\theta_f) L^{(\beta-1)} - R = 0 \qquad or \tag{7.3}$$

$$L_f = \left[\frac{\beta p_k V(\theta_f)}{R} \right]^{\frac{1}{1-\beta}}$$

As expected, loan demand is decreasing in the loan rate. Differentiating (3),

$$\frac{\partial L_f}{\partial R} = -\frac{1}{(1-\beta)} \frac{L_f}{R} < 0 \tag{7.4}$$

2. Loan applicants with higher management quality demand larger loans.

Inspection of (3) shows that at any R, firms with higher θ demand higher loans.

3. In $<R,L>$ space, the isoprofit curve of the loan applicant is strictly quasi-concave in L, with zero slope at the point of loan demand.

The slope of a firm's isoprofit curve is given by[57]:

$$\frac{dR}{dL}\big|_{\phi_f} = \frac{\beta p_k V(\theta_f) L^{\beta-1} - R}{L} \tag{7.5}$$

$$\gtreqless 0 \ \ as \ \ \beta p_k V(\theta_f) L^{\beta-1} - R \gtreqless 0$$

Thus, in $<R,L>$ space with L on the horizontal axis, a firm's isoprofit curve has positive slope where the value of the marginal productivity of a loan exceeds the loan rate, zero slope at the point where the loan demand schedule intersects the isoprofit curves, and negative slope at loan levels beyond this. Further, differentiating (5) with respect to L:

$$\frac{\partial}{\partial L}\left[\frac{dR}{dL}\big|_{\phi_f}\right] = \frac{(\beta-1)\beta p_k V(\theta_f) L^{\beta-2} - \frac{dR}{dL}}{L} \tag{7.6}$$

$(\beta-1) < 0$ implies that whenever $\frac{dR}{dL} = 0$, $\frac{\partial}{\partial L}\left[\frac{dR}{dL}\big|_{\phi_f}\right] < 0$ so that the borrower's isoprofit curve is strictly quasi-concave in L.

4. Borrowers' isoprofit curves satisfy the single-crossing property.

The isoprofit curves for these firms satisfy the single-crossing property required for screening to be feasible. Differentiating (5) with respect to θ:

$$\frac{\partial}{\partial \theta}\left[\frac{dR}{dL}\big|_{\phi_f}\right] = \beta p_k V'(\theta_f) L^{\beta-2} > 0 \tag{7.7}$$

Firms with lower θ have less steep isoprofit curves. Thus at loan sizes below their profit-maximizing levels, where the isoprofit curves are positively sloped, they are willing to accept a smaller marginal decrease in the loan rate for a marginal decline in loan size. In this range of their isoprofit curves, the marginal productivity of a loan exceeds its marginal cost for both borrowers. But θ_2's higher productivity means that he values the marginal loan more and must be

---

[57] The notation $\frac{dx}{dy}\big|_{\phi_f}$ indicates that the derivative is taken along the isoprofit contour of type f.

compensated for a decrease by a larger decrease in loan rate.

BANKS

A bank's profit function is assumed to be additively separable in the loans extended to different firms. Expected profit for loans to an identified firm is given by:

$$E\pi = q_f R_f L_f - I L_f \tag{7.8}$$

where $I = (1+i)$ is the market deposit rate factor.

The bank isoprofit curve in $< R, L >$ space has slope

$$\frac{dR}{dL}\big|_{\pi} = \frac{I - q_f R_f}{L_f} \tag{7.9}$$

The bank breaks even on loans to type θ_f at $R_f = I / q_f$ so that, given I, for each type there is an isoprofit schedule, with zero slope, giving the lowest loan rate the bank can charge type θ_f for nonnegative profits.

If the bank is unable to distinguish between types, its profit function is given by:

$$E\pi = (q_1\gamma_{k1} + q_2\gamma_{k2})RL - I \gamma_k L \tag{7.10}$$

where γ_k = the proportion of agents in the y_k industry, γ_{kf} = proportion of agents of type θ_f, $f = 1,2$ in the y_k industry, and L = loan size accorded to indistinguishable y_k firms. The isoprofit curve has slope:

$$\frac{dR}{dL}\big|_{\pi} = \frac{\gamma_k I - (q_1\gamma_{k1} + q_2\gamma_{k2})R}{(q_1\gamma_{k1} + q_2\gamma_{k2})L} \tag{7.11}$$

Break-even on these pooling contracts occurs where the isoprofit curve has zero slope in $< R, L >$ space at

$$R_p = \frac{\gamma_k I}{(q_1\gamma_{k1} + q_2\gamma_{k2})} \tag{7.12}$$

where R_p is the break-even pooling loan rate.

THE MARKET

The analytical framework introduced by Rothschild-Stiglitz-Wilson (RSW) ([1976] and [1977], respectively) describes behaviour in markets with asymmetric information as character-

ized by contract-setting, where the contracts specify both the price and quantity (as well, possibly, as other sorting variables). RSW implicitly modelled the market process as a two-stage game where (in the context of the market considered here) banks offer contracts in the first stage and the potential borrowers select among the available contracts in the second stage (see Hellwig [1987]).

Symmetric Information

In order to provide a benchmark against which to compare the results obtained under asymmetric information, the symmetric information case is considered first. Each bank maximizes its separable profit function subject only to the individual rationality constraints that each borrower obtains at least the equilibrium profit.

$$
\begin{aligned}
\underset{R_1, L_1, R_2, R_2}{Max} \quad & \gamma_{k1}[q_1 R_1 L_1 - I L_1] + \gamma_{k2}[q_2 R_2 L_2 - I L_2] \\
& + \lambda_1 \left[q_1 p_k L_1^\beta V(\theta_1) - q_1 R_1 L_1 - \hat{\phi}_1 \right] \\
& + \lambda_2 \left[q_2 p_k L_2^\beta V(\theta_2) - q_2 R_2 L_2 - \hat{\phi}_2 \right]
\end{aligned}
$$

where $\hat{\phi}_f$ = equilibrium profit for firm of type θ_f, λ_f = multiplier for individual rationality constraint for firm of type θ_f, L_f^s denotes the loan size when information is symmetric, and $\beta L_f^{\beta-1} V(\theta_f)$ is the marginal productivity of loan size L_f. The first order conditions give:

$$
\beta \left[L_f^s \right]^{\beta-1} V(\theta_f) = \frac{I/q_f}{p^k}, \qquad f = 1,2 \tag{7.13}
$$

Since equilibrium with perfect information and perfect competition entails equality between the marginal productivity of a factor and the ratio between its price and the output price, the price of L_f is $R_f = I / q_f$. From (7.13) the symmetric equilibrium loan sizes are obtained as:

$$
L_1^s = \left[\frac{p_k \beta V(\theta_1)}{R_1} \right]^{\frac{1}{1-\beta}} \quad and \tag{7.14i}
$$

$$
L_2^s = \left[\frac{p_k \beta V(\theta_2)}{R_2} \right]^{\frac{1}{1-\beta}} \tag{7.14ii}
$$

Thus, with symmetric information both types receive the loan demanded at their respective loan rates, as can be confirmed by comparing (7.3) and (7.14).

In this complete information setting, this result can be viewed as reflecting banks' profit maximization in separate submarkets for each type. In each submarket it can be assumed that each bank treats the rate of interest as parametric and chooses loan size[58]. When the assumption of complete information is relaxed the hypothesis that clients are sorted requires that banks must choose the loan rate and loan size (the contracts) required for sorting. Explicit consideration of strategic behaviour becomes necessary. Adoption of a game-theoretic approach in even the present symmetric information environment allows for consistent treatment. We may view the offer and acceptance of contracts as a game with sequential strategies and look for a subgame perfect equilibrium (Selten [1975]) as is appropriate in a repeated game. The economy deposit rate factor is I and all banks compete by making separate offers to the two types in the first stage, the offers being accepted or rejected in the second. Consider the second stage when borrowers choose among the contracts available to them. Given any rate of interest, R, a contract specifying a loan size $L \neq [R^{-1}(\beta p_k V(\theta_f))]^{1/(1-\beta)}$ will be rejected by borrowers of type θ_f if a contract with L as defined by (7.3) is available since the latter lies on their demand curve. Thus at R only contracts with L as defined by (7.3) cannot be upset by alternative offers. This holds for every R. No contract with $R < I/q_f$ will be on offer since, from (7.8), this implies negative profits. Given these two properties which the contracts on offer in stage 2 must satisfy in order to be Nash equilibria for that stage, we consider contract offers in stage 1. At stage 1 $R \geq I/q_f$ but Bertrand competition implies $R = I/q_f = R_f$. Individual profit maximization implies that with R_f L is given by (7.13). But these are precisely the contracts that cannot be upset at stage 2, for any contract off the demand curve of type θ_f there is an incentive to offer a contract which will attract all borrowers of that type in stage 2. Thus contracts defined by (7.14) are a subgame perfect equilibrium.

Under complete information the relative loan sizes obtained by the two types depend on the assumption made about expected returns. Loan sizes are allocated according to the expected return of the debtor and coincide with the loan size they demand (see Figure 7.1 which illustrates

---

[58] Compare Gale's [1987b] treatment in the asymmetric information case. The assumption of a contract per market and ex post matching for price quotes allowed a Walrasian equilibrium.

how loan size depends on expected return, and note the same relationship in (7.14)). The higher

loan rate charged to θ_2 compensates the bank for the lower probability of repayment. Thus banks'

marginal rates of substitution (MRS) between R_f and L_f equals that of the type θ_f agents in each

market and the solutions to the banks' and borrowers' individual maximization problems coin-

cides with the outcome required for efficient production in the economy. With asymmetric infor-

mation, however, this coincidence of ends is no longer automatically available.

Figure 7.1: *The Symmetric Information Loan Contracts*

Loan size depends on expected return: taking $V(\theta_1)$, $V(\theta_2)$ and q_2 as fixed and letting q_1 vary, at low R_1, R_{1L}, $q_1V(\theta_1) > q_2V(\theta_2)$, and $L_1^s > L_2^s$, at R_{1M} $q_1V(\theta_1) = q_2V(\theta_2)$ and $L_1^s = L_2^s$, at R_{1H} $q_1V(\theta_1) < q_2V(\theta_2)$ and $L_1^s < L_2^s$.
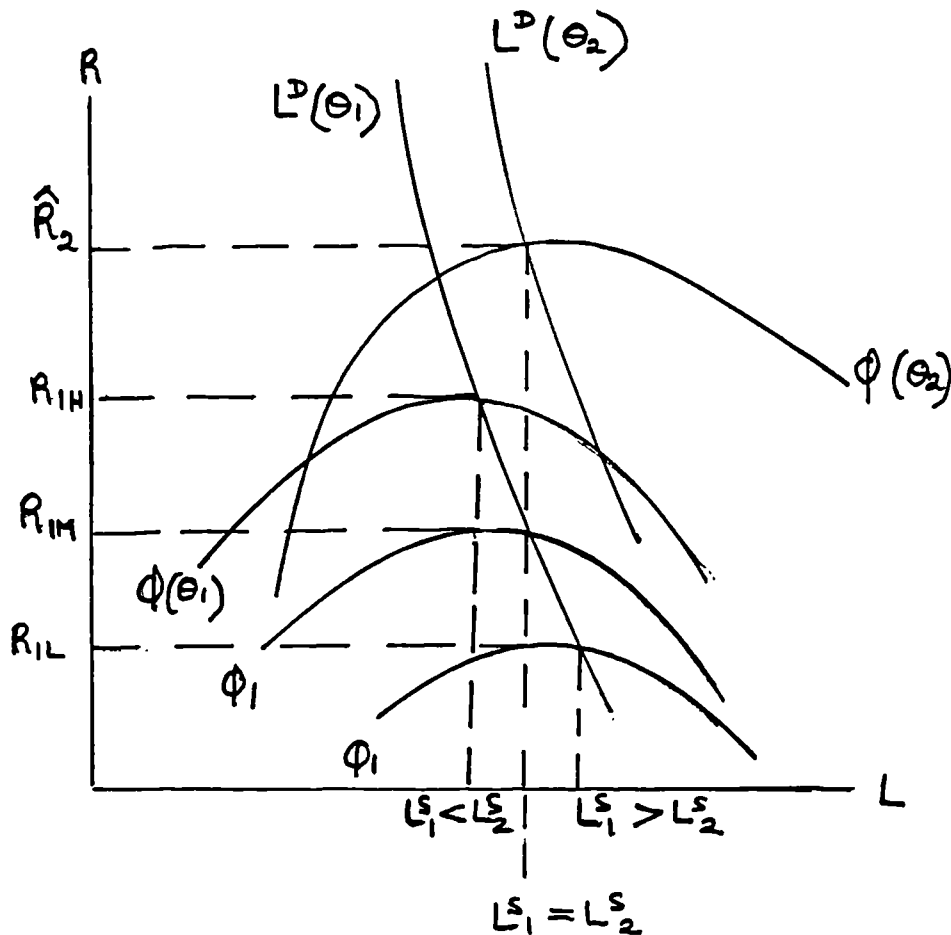
Asymmetric Information

We now assume that banks know the distribution of types on the market and their success probabilities and qualities but cannot identify the types. The symmetric information contracts make banks vulnerable to losses which occur on any loan offered to θ_f at $R < I/q_f$ (see (7.9)). The first two points below indicate the constraints on a bank's options and the third how screening contracts address these.

1. First best contracts are not incentive compatible.

The lower loan rate set for θ_1 induce θ_2 types to claim to be θ_1, that is, the first best contracts are not incentive compatible. Feasibility of first best contracts when types are private information requires that θ_2's profit at the symmetric information contract $< R_2^s, L_2^s >$ be at least as great as the profit obtained from pretending to be θ_1. The following weak inequality must be satisfied:

$$q_2 [p_k (L_2^s)^\beta V(\theta_2) - R_2^s L_2^s] \geq q_2 [p_k (L_1^s)^\beta V(\theta_2) - R_1^s L_1^s]$$

Substituting for L_1^s, L_2^s, R_1^s and R_2^s and simplifying, this requires :

$$\left[q_2 V(\theta_2) \right]^{\frac{\beta}{1-\beta}} \left[\frac{1}{\beta} - 1 \right] \geq \left[q_1 V(\theta_1) \right]^{\frac{\beta}{1-\beta}} \left[\frac{1}{\beta} - \frac{V(\theta_1)}{V(\theta_2)} \right]$$

Since, with $\frac{1}{\beta} \geq 1$ and $V(\theta_1) < V(\theta_2)$

$$\left[\frac{1}{\beta} - 1 \right] < \left[\frac{1}{\beta} - \frac{V(\theta_1)}{V(\theta_2)} \right]$$

this inequality will clearly be violated if either, a) the θ_2 types' return is a mean preserving spread of the θ_1 types' return, i.e. $q_2 V(\theta_2) = q_1 V(\theta_1)$ and $q_2 < q_1$, or b) the expected output of θ_1 types exceeds that of θ_2 types, $q_1 V(\theta_1) > q_2 V(\theta_2)$. If $q_2 V(\theta_2) > q_1 V(\theta_1)$, the inequality will not hold if

$$\left[\frac{1 - \beta V(\theta_1)/V(\theta_2)}{1 - \beta} \right] > \left[\frac{q_2 V(\theta_2)}{q_1 V(\theta_1)} \right]^{\frac{\beta}{1-\beta}}$$

Inspection of Figure 7.1 indicates that this last possibility occurs when, given $V(\theta_1)/V(\theta_2)$, q_1/q_2 is so low that R_1 is sufficiently close to R_2 that θ_2 types prefer their larger loan sizes at a higher rate than θ_1's lower loan size and only slightly lower loan rate. Asymmetric information then creates no identification problem for a bank and we exclude this possibility by assumption.

Given this assumption, Figure 7.1 shows that θ_2 can move to higher levels of profit by claiming to be θ_1. But the lending bank loses on any loans made to type θ_2 at a loan rate below $\hat{R}_2 = I / q_2$. In the absence of complete information, a bank can i) offer contracts that sort customers, ii) offer a single contract at a loan rate reflecting the average probability of success, iii) refuse loans to all borrowers in the y_k sector. The last possibilities give rise to Stiglitz-Weiss credit rationing since, we note:

2. Raising the loan rate in a pooling contract produces adverse selection.

Consider the present model with many types. Suppose banks are unable to distinguish these types, but set a single rate of interest and lend each applicant the loan demanded at that rate of interest. No firm for whom $q [p_k L^\beta V(\theta) - RL] < 0$ will apply for a loan. Let the values for which this condition is satisfied with equality be:

$$\frac{\overline{L}^\beta V(\overline{\theta})}{\overline{L}} = \frac{R}{p^k} \quad \textit{that is}$$

$$\overline{AP}_L = \frac{R}{p^k}$$

where $\overline{AP}_L$ is the average loan product for the marginal type, $\overline{\theta}$. Any θ with $AP_L < \overline{AP}_L$ will not apply for a loan. But as R increases the cutoff value of AP increases. (For any θ an increase in R reduces profit in the case of success so that expected profit decreases). Firms with lower AP (lower θ) will leave the market. But firms with lower AP are also less risky and hence more desirable from the bank's point of view.

3. Screening contracts permit separation

Instead of the adverse selection effect, it is assumed that banks are sufficiently knowledgeable of profitable market opportunities to screen. We conjecture that equilibrium is characterized by bank offers of a set of contracts among which borrowers choose according to their type. In the conjectured equilibrium, the contract variables are solutions to the bank constrained profit maximization problem. These constraints are the individual rationality constraints and the following incentive compatibility constraints:

$$q_1 [p_k L_1^\beta V(\theta_1) - R_1 L_1] \geq q_1 [p_k L_2^\beta V(\theta_1) - R_2 L_2] \tag{7.15i}$$

$$q_2 [p_k L_2^\beta V(\theta_2) - R_2 L_2] \geq q_2 [p_k L_1^\beta V(\theta_2) - R_1 L_1] \tag{7.15ii}$$

Constraint (7.15i) requires that the LR type, θ_1 prefer the contract intended for him to that intended for θ_2, and conversely for constraint (7.15ii). It is conjectured that if constraint (7.15ii) is binding, constraint (7.15i) will not be[59]. The Lagrangean for the bank's problem can be written:

$$\underset{R_1,L_1,R_2,L_2}{Max} \quad \gamma_{k1}[q_1R_1L_1 - IL_1] + \gamma_{k2}[q_2R_2L_2 - IL_2] \tag{7.16}$$

$$+ \eta \left\{ q_2 p_k L_2^\beta V(\theta_2) - q_2 R_2 L_2 - [q_2 p_k L_1^\beta V(\theta_2) - q_2 R_1 L_1] \right\}$$

$$+ \lambda_1 \left\{ q_1 p_k L_1^\beta V(\theta_1) - q_1 R_1 L_1 - \bar{\phi}_1 \right\}$$

$$+ \lambda_2 \left\{ q_2 p_k L_2^\beta V(\theta_2) - q_2 R_2 L_2 - \bar{\phi}_2 \right\}$$

where η is the Lagrange multiplier for incentive compatibility constraint (7.15ii), λ_1, λ_2 are the multipliers for θ_1's and θ_2's individual rationality constraints, and $\bar{\phi}_f$ = the profit level type θ_f would expect to obtain in a competitive equilibrium, i.e. their reservation profit levels. L_f^c and R_f^c will denote the loan size and loan rate received by type θ_f in the screening contract. The first order conditions give equations (7.17) to (7.20). The equalities follow from the fact that distinct contracts must be offered in order for screening to be effective, and a positive loan size will obviously bear a positive rate of interest, i.e. L_1 and L_2 must both be positive, giving the equalities in (7.18) and (7.20), and hence so must R_1 and R_2, giving (7.17) and (7.19).

$$R_1: \quad \gamma_{k1}q_1 + \eta q_2 - \lambda_1 q_1 = 0 \tag{7.17}$$

$$L_1: \quad \gamma_{k1}[q_1R_1^c - I] - \eta q_2[p_k\beta(L_1^c)^{\beta-1}V(\theta_2) - R_1^c] + \lambda_1 q_1[\beta p_k(L_1^c)^{\beta-1} - V(\theta_1) - R_1^c] = 0 \tag{7.18}$$

$$R_2: \quad \gamma_{k2} - \eta - \lambda_2 = 0 \tag{7.19}$$

$$L_2: \quad \gamma_{k2}[q_2R_2^c - I] + \eta q_2[p_k\beta(L_2^c)^{\beta-1}V(\theta_2) - R_2^c] + \lambda_2 q_2[\beta p_k(L_2^c)^{\beta-1}V(\theta_2) - R_2^c] = 0 \tag{7.20}$$

Using (7.17) and (7.18) to eliminate λ_1, simplifying and rearranging, the following value is obtained for multiplier η:

[59] We note that i) incentive compatibility requires that at least one self-selection constraint binds, ii) that at most one incentive constraint binds: assume that both bind and add the resulting equations; this implies that $V(\theta_1) = V(\theta_2)$, that the types do not differ in productivity, contrary to assumption. iii) We can also check that if (7.15i), the constraint on θ_1, is binding, then (7.15ii) does not. Equality in (7.15i) implies that the difference between the LHS and RHS of (7.15ii) is $p_k[L_1^\beta - L_2^\beta][V(\theta_1) - V(\theta_2)]$. But $V(\theta_1) < V(\theta_2)$ and the nonnegativity of the difference required for self-selection implies that $L_1^\beta < L_2^\beta$ - as we will see is the case in the screening contract. But the implied contracts have θ_2 signalling by a loan size above his perfect information choice, while θ_1 receives his first best contract. But this violates the single crossing property since θ_2 has a higher marginal cost of signalling. In short such a contract would not be incentive compatible.

$$\eta = \frac{\gamma_{k1} q_1 [\beta p_k (L_1^c)^{\beta-1} V(\theta_1) - I/q_1]}{q_2 \beta p_k (L_1^c)^{\beta-1} [V(\theta_2) - V(\theta_1)]} \tag{7.21}$$

Substituting (7.19) in (7.20) gives:

$$\beta p_k (L_2^c)^{\beta-1} V(\theta_2) - \frac{I}{q_2} = 0 \tag{7.22}$$

Bertrand competition ensures that at any $R_f > I/q_f$, the loan rate will be bid down to equality. Equation (7.22) therefore indicates that in the contract for θ_2 loan size is set such that the marginal productivity of the loan is equated to the ratio of loan rate (factor price) and the price of the good, as in the full information case. If the contract intended for θ_1 specified a similarly-determined loan size, the numerator of equation (7.21) would be zero. θ_1 would also be receiving his first best contract. But our consideration of θ_2's profits evaluated at $< R^s, L^s >$ in point 1 above shows that it is not incentive compatible for both types to receive their first best contracts. Hence

$$\beta p_k (L_1^c)^{\beta-1} V(\theta_1) - \frac{I}{q_1} > 0 \tag{7.23}$$

therefore $\eta > 0$ and constraint (7.15ii) is binding.

Thus, in the conjectured screening equilibrium, the following contracts will be offered. Equation (7.22) gives the contract for type θ_2 as:

$$< R_2^c, L_2^c >: \quad R_2^c = R_2^s = R_2 = \frac{I}{q_2} \tag{7.14iibis}$$

$$L_2^c = L_2^s = \left[\frac{p_k \beta V(\theta_2)}{R_2} \right]^{\frac{1}{1-\beta}}$$

As usual in screening models, the 'lowest' quality types, who have nothing to gain from being identified, receive the first best contract. Quality here is judged from the point of view of the uninformed, namely the bank, who prefers low risk borrowers. Substituting L_2^c from (7.14ii) into (7.15ii) as an equality gives the contract for type θ_1 as[60]:

$$< R_1^c, L_1^c >: \quad R_1^c = R_1^s = R_1 = \frac{I}{q_1} \tag{7.24}$$

$$L_1^c = \frac{p_k [y_k(L_1^c, \theta_2) - (1-\beta) y_k(L_2^c, \theta_2)]}{R_1}$$

---

[60] We note that L_1^c depends on θ_2, the Greenwald-Stiglitz externality discussed in Chapter 8.

where $y_k(L,\theta)$ = output of y_k evaluated at stated values. Recall from (7.5) and (7.6) that the isoprofit curves are strictly quasi-concave in L with positive slopes at L input levels below the loan demand curve. The expression in (7.23) is the slope of the θ_1 isoprofit curve at the screening contract - compare (7.5). Therefore positive (7.23) (positive η) indicates that θ_1 types receive a loan size smaller than demanded at R_1. θ_1 types, who have a lower marginal cost of signalling, self-select by their willingness to accept a smaller than optimal loan. We can now state the following.

Proposition: $L_2^c > L_1^c$ i.e. the loan size in the screening contract selected by type θ_2 exceeds the loan size in the screening contract chosen by type θ_1.

We may first consider a diagrammatic demonstration of this. Consider the set of contract offers faced by θ_1 and θ_2. The contract $<R_2,L_2^c>$ is optimal for θ_2. Therefore, in Figure 7.2 the other contracts on offer must lie in the vertically shaded region. Since profits, ϕ, are strictly decreasing in R, we know that θ_1 prefers any rate of interest below R_2. We also know that loans to θ_1 are made at the breakeven rate R_1. Therefore the alternative offers which θ_1 prefers must lie in the horizontally striped region in Figure 7.2. The single crossing property says that at any $<R,L>$, θ_2 has a steeper isoprofit curve, and these two isoprofit curves cannot cross again. Then the intersection of the two shaded regions (where there is a contract preferred by θ_1 but not by θ_2) must lie to the left of L_2^c and below R_1.

Figure 7.2: *Screening by low-risk acceptance of a smaller loan*
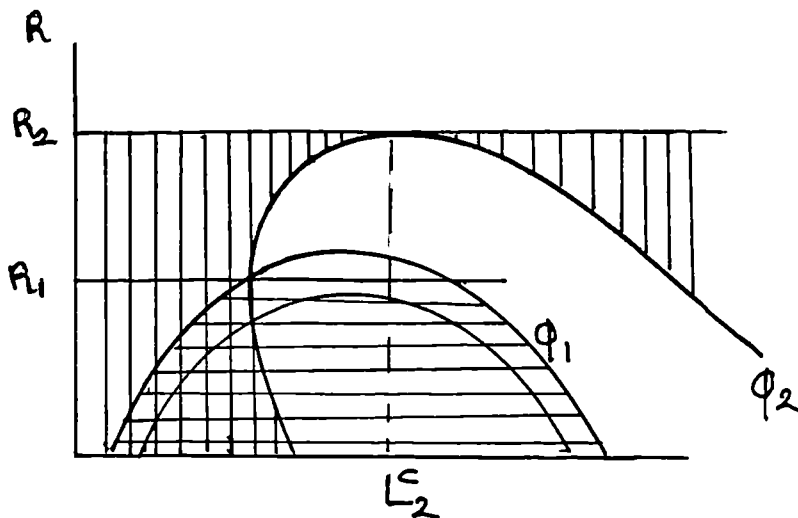
We may also proceed by contradiction. Assume $L_1^c > L_2^c$. From equation (7.14ii), $L_2^c = L_2^s$. Therefore, $L_1^c > L_2^s$. Consider the binding constraint:

$$[p_k (L_2^c)^\beta V(\theta_2) - R_2 L_2^c] = [p_k (L_1^c)^\beta V(\theta_2) - R_1 L_1^c]$$

The RHS is increasing in L since

$$\frac{\partial [p_k (L_1^c)^\beta V(\theta_2) - R_1 L_1^c]}{\partial L} = \beta p_k (L_1^c)^{\beta-1} V(\theta_2) - R_1 > 0$$

The inequality follows from (7.23) and $V(\theta_2) > V(\theta_1)$. Therefore,

$$p_k (L_2^c)^\beta V(\theta_2) - R_2 L_2^c > p_k (L_2^s)^\beta V(\theta_2) - R_1 L_2^s$$

But this implies that θ_2 prefers higher interest rates, since $R_2 > R_1$ - a contradiction. Hence $L_1^c \leq L_2^c$, but, with $R_1 < R_2$, the constraint requires that the strict inequality must hold.

Thus, in order to screen, banks cannot allocate financing according to expected return: the LR always receive the smaller loan, irrespective of expected return: as noted earlier, LR expected return may be smaller, equal to, or larger than HR.

Proposition: A binding incentive compatibility constraint (7.15ii) for θ_2 implies that the constraint for θ_1 does not bind.

From (17.15ii), the binding constraint we have:

$$[p_k (L_2^c)^\beta V(\theta_2) - R_2 L_2^c] = [p_k (L_1^c)^\beta V(\theta_2) - R_1 L_1^c]$$

Adding and subtracting $p_k (L_1^c)^\beta V(\theta_1)$ on the RHS:

$$[p_k (L_2^c)^\beta V(\theta_2) - R_2 L_2^c] = [p_k (L_1^c)^\beta V(\theta_2) - R_1 L_1^c] + p_k (L_1^c)^\beta V(\theta_1) - p_k (L_1^c)^\beta V(\theta_1) \quad (7.25)$$

Adding and subtracting $p_k (L_2^c)^\beta V(\theta_1)$ on the LHS:

$$[p_k (L_2^c)^\beta V(\theta_2) - R_2 L_2^c] = [p_k (L_2^c)^\beta V(\theta_2) - R_2 L_2^c] + p_k (L_2^c)^\beta V(\theta_1) - p_k (L_2^c)^\beta V(\theta_1) \quad (7.26)$$

Equating (7.25) and (7.26) and rearranging we obtain on the LHS the incentive compatibility constraint for θ_1:

$$p_k (L_1^c)^\beta V(\theta_1) - R_1 L_1^c - [p_k (L_2^c)^\beta V(\theta_1) - R_2 L_2^c]$$
$$= p_k [V(\theta_2) - V(\theta_1)][(L_2^c)^\beta - (L_1^c)^\beta] > 0 \quad (7.27)$$

$V(\theta_2) > V(\theta_1)$ and $L_2^c > L_1^c$ implies that the RHS of (7.27) is positive.

We now show that these separating contracts are a sequential equilibrium.

Screening contracts as a sequential equilibrium.

We argue that the conjectured screening contracts are an equilibrium if either or, *a fortiori*, both of the following assumptions hold:

Assumption 1: The proportion of θ_1 types in the y_k industry population is small .

Let $\gamma_{k1}/\gamma_{k2} = \gamma_1$ be the population share of types $\theta1$ in industry y_k. The average success probability in the industry can be written $\bar{q} = \gamma_1 q_1 + (1-\gamma_1) q_2$ so that the assumption of small γ_1 implies that the average success probability is near to the lower success probability of θ_2. While this assumption (or the next) is necessary for the equilibrium we propose, it is in fact the McKinnon-Shaw assumption of many high productivity, high risk activities in the non-traditional sectors of the lagging economies.

Assumption 2: The difference $\left[\dfrac{dR}{dL} \big|_{\bar{q}_1} - \dfrac{dR}{dL} \big|_{\bar{q}_1} \right]$ is large.

The assumption is that the θ_2 isoprofit curve is quite steeply sloped relative to that of θ_1 or, equivalently, that θ_1's marginal cost of signalling is quite low relative that of θ_2. We have already shown - see (7.5) - that the low risk have lower marginal cost and that this is required for incentive compatibility.

As discussed we may view the contract signing on the credit market as a sequential game in which banks offer contracts in stage 1 and borrowers choose among the contracts on offer in stage 2. Given the sequential nature of the game and the private information about borrower types, it is appropriate to use an equilibrium concept that ensures credibility when information is imperfect. We therefore look for a sequential equilibrium (SE) (Kreps and Wilson [1982]), following Hellwig's [1987] method. A strategy combination is a sequential equilibrium if there are consistent beliefs such that each player's strategy prescribes at every information set a choice which is optimal with respect to those beliefs. In the first stage banks know the distribution of types in the sector and their preferences but cannot identify the types. Each bank offers a set of contracts which specify an interest rate and a loan size. They have beliefs about the contracts other banks

will offer given their common knowledge of the structure of the game and borrowers' strategies in the second stage. SE requires that they offer contracts that are a best response to the offers of other banks and to the decisions they expect borrowers to make at stage 2. In the second stage, since the informed know the contracts placed on the market in the first stage, SE requires only that their strategies be sequentially rational, that is, that a borrower chooses the contract that is optimal for him among those available, regardless of the contracts offered at stage 1.

Consider this second stage choice of the borrowers. With the conjectured screening contracts available, type θ_1 will choose $<R_1^c, L_1^c>$, since by (7.27)

$$\phi_{\theta_1}\left[R_1^c, L_1^c\right] > \phi_{\theta_1}\left[R_2^c, L_2^c\right]$$

Type θ_2 will choose $<R_2^c, L_2^c>$, since

$$\phi_{\theta_2}\left[R_2^c, L_2^c\right] = \phi_{\theta_2}\left[R_1^c, L_1^c\right]$$

by binding constraint (7.15ii). He is indifferent between $<R_2^c, L_2^c>$ and $<R_1^c, L_1^c>$ but the former lies closer to his demand curve.

Suppose a pooling contract were also available. This contract will always be chosen by θ_2 in preference to his screening contract because $\bar{q} > q_2$ means that $R_2 > R_p$, where R_p is the pooling contract loan rate. However, if only θ_2 chooses the pooling contract, the bank offering it will make losses because R_p does not compensate the bank for the risk borne with θ_2. Therefore the deviation to the pooling equilibrium offer in the first stage is only profitable if θ_1 types also demand the contract.

In choosing $<R_1^c, L_1^c> \theta_1$ bears a cost in terms of lower loan size than preferred at R_1. However, he thus avoids the subsidy he would have to pay for θ_2 in the pooling contract. Therefore θ_1 does not take the pooling contract if the cost of doing so in terms of loan rate subsidy outweighs the gain through loan size. There is no net gain from the pooling contract if R_p is high. R_p is high if the average success probability required by the bank is low, that is, if γ_1 is low, as in Assumption 1. This is illustrated in Figure 7.3. With the screening contracts, θ_1 is on isoprofit curve ϕ_1'. As long as ϕ_1' lies below R_p, θ_1 finds the screening contract more profitable. Only if the pooling rate is as low as $\bar{R}_p$ would he find the pooling contract attractive, but we have excluded this by

assumption.

**Figure 7.3:** *The Screening Contracts with Small* γ_1
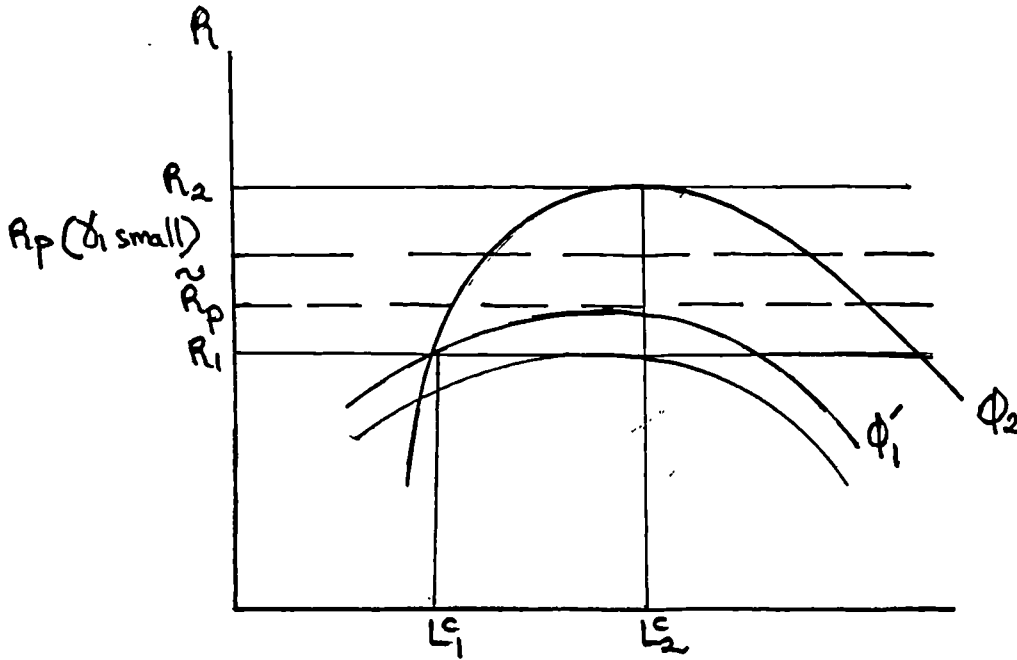
The extent of the gain via loan size in the pooling contract also depends on the reduction in loan size which the screening contract must include in order to sort. This depends on the relative magnitudes of the types' MRS. If type θ_1's marginal cost of signalling is very low relative to θ_2's, that is, if θ_1's isoprofit curve is relatively shallower, the bank can sort with relatively lower reductions in loan size relative to first best. In Figure 7.4, with isoprofit ϕ_2^H the bank has only to reduce loan size by $d1 < d2$ which is the reduction required with isoprofit ϕ_2^L, which represents a relatively lower marginal cost of being screened for θ_1. Thus under Assumption 2 θ_1 will not accept a pooling contract.

Under either assumptions 1 or 2 sequential rationality implies that θ_2 chooses a screening contract at stage 2. Given this choice a pooling contract will be strictly dominated by the screen-

ing contract. Therefore in stage 1 the only contracts offered will be screening contracts. All

banks offer these screening contracts and borrowers are allocated randomly among them in stage

2. Banks' beliefs regarding the strategy of borrowers and other banks are proved consistent and

the screening contracts are a sequential equilibrium.

Figure 7.4: *The Screening Contracts with Large*
$MRS(\theta_2)/MRS(\theta_1)$
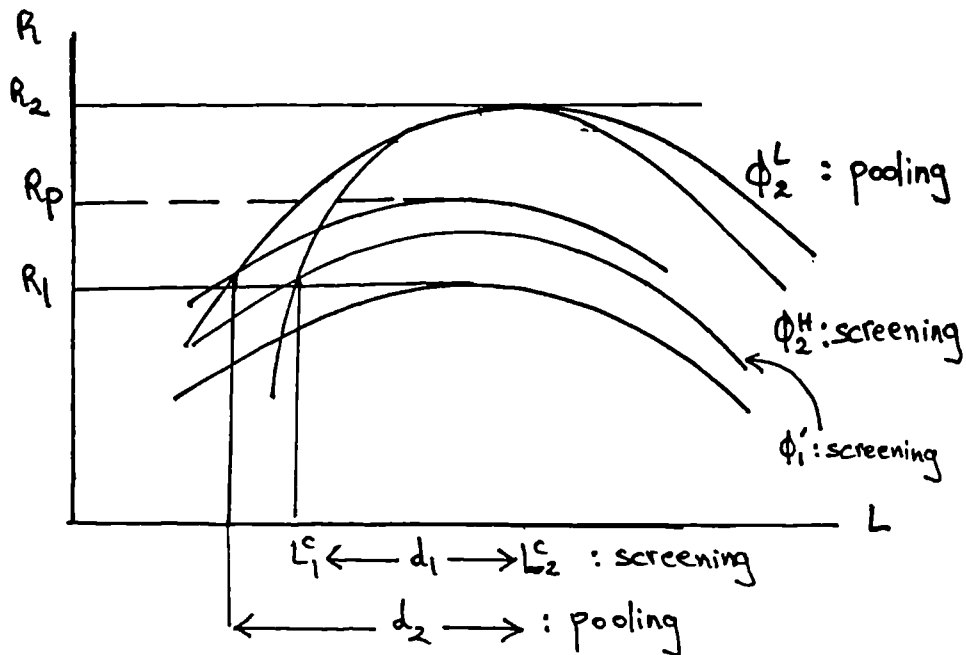
The assumption above on the relative marginal costs of being screened is that derived by

Riley [1985] to ensure existence of a Nash equilibrium in a general screening model. That about

relative population shares was first discussed in the Rothschild-Stiglitz [1976] paper. Hellwig's,

op. cit., discussion of the two-stage Rothschild-Stiglitz-Wilson formulation breaks the population

share assumption in arguing that that game has no SE. Hellwig adds a third stage in which the

uninformed banks reject or accept the loan applications made in the second stage of the R-S-W

game. With the 3-stage structure the screening contracts we have argued are accepted at stage 2

are no longer an SE because, in the absence of our assumptions, it is profitable to deviate to pool-

ing contract. The pooling contract can always be in turn upset by a separating contract which

attract the LR, leading to losses on the pooling contract (the original R-S-W argument that

pooling is never an equilibrium). The two-stage game therefore has no equilibrium in pure strategies. Hellwig argues that in the 3-stage game the pooling contract cannot be upset by a separating contract and is the most plausible SE. Deviation to the separating contract is unprofitable because a demand for the contract will be rejected by the banks at stage 3 and it will therefore not be demanded in stage 2[61].

Hellwig's model has the important advantage of generality: there is no need to make a specific assumption about population share or relative MRS in order to obtain an equilibrium. Consequently, the Stiglitz and Weiss original argument that asymmetric information leads to rationing receives strong support. This is reinforced by the rationing results found in more general models of the debt contract (Gale and Hellwig, op. cit., and Williamson, op. cit.). We have not chosen to use this 3-stage formulation because we wish to consider the prescriptive implications of assuming that banks' forte lies in making arrangements that produce reliable information about borrowers' investment opportunities, as argued by McKinnon and Shaw. The use of a formulation that produces the rationing result does not allow banks the most favourable informational role. Further, as argued, the first assumption is consistent with Shaw's view of the LDC environment.

7.5: Interim Conclusion

Our model has accomplished three tasks. First, we have shown that in models of screening by loan size a sequential equilibrium can be obtained with assumptions similar to those used by authors such as Milde and Riley. Since screening models are based on the premise that banks offer contracts among which borrowers choose, they are two-stage asymmetric information games in which the sequential equilibrium is the appropriate equilibrium concept. Second, we have demonstrated that while banks are able to generate information which sorts risk classes, as

---

[61] Let the pooling contract be PC and the PC equilibrium-breaking separating contract be SC. If SC applications are accepted in stage 3, the accepting bank must believe that its applicants are LR (that its applicants have an above-average success probability) therefore banks offering the PC must have below-average applicants. But then PC will make losses and applications for it will not be accepted at stage 3. But this implies that the applicants for SC are the entire risk population and SC will experience losses. There is therefore a contradiction to the assumption that SC is accepted.

hypothesized by McKinnon and Shaw, the resulting loans are not allocated on the basis of expected social return as would occur under symmetric information. In order to distinguish themselves for a lower loan rate, the low risk must always accept a lower loan size. Although all borrowers obtain loans, there is no assurance that the highest productivity projects will be fully financed. Third, our explicit derivation of the contract variables serve to identify the maximized profit functions and demand functions of the sector which we assume banks have to screen. We use these functions in Chapter 8 to show that the competitive equilibrium with screening is not constrained Pareto efficient even though banks are sorting perfectly. Incentive compatibility constrains the screened to be off their optimum and therefore the envelope theorem does not hold for their indirect profit functions so that price changes can have allocative effects.

Chapter 8: Credit Market Policy when Banks Screen: A General Equilibrium Approach

8.1: Introduction

The key feature of credit markets with asymmetric information is the heterogeneity of borrowers whose quality and success probabilities cannot be identified *a priori*. However, the price they are willing to pay for a given loan depends on their type. When price changes therefore demand changes both quantitatively and qualitatively. Banks acquire commodities - loan accounts - whose value they cannot assess. Chapter 7 discussed and illustrated how banks can cope with this risk by rationing or by offering screening contracts.

Previous models of credit market policy have considered government intervention when contracts specify only the loan rate. It is well-known that loan-rate-only contracts lead to credit rationing and/or underinvestment or to overinvestment. However, until Greenwald and Stiglitz's [1986] investigation of the Pareto efficiency of economies with incomplete markets and imperfect information it has been widely assumed that separating equilibria are constrained Pareto efficient. Therefore, the Pareto efficiency of credit markets when banks sort perfectly has not, it is thought, previously been examined. This chapter adopts the Greenwald and Stiglitz methodology, incorporating Chapter 7's screening contracts into a modified version of their general equilibrium model. Given the credit market conditions of Chapter 7's model, we find that the appropriate policy is to subsidize the highest interest rates. This result can be explained as follows.

Screening is achieved in our model through the low risk's acceptance of a loan size below the first best level. There is less investment, with a lower average success probability, being financed than in the first best. The implicit policy objective is to increase the expected return of aggregate investment, without disturbing the self-selection induced by the intervention free competitive equilibrium contracts. High risk (HR) indifference to low risk (LR) contracts should be maintained in order that banks continue to screen profitably. While a subsidy on higher interest rates increases the loan size taken by the HR it permits the contract offered for the LR to specify a sufficiently larger loan size to improve the risk mix of financed investment, simultaneously rais-

ing total financing. The HR's subsidy maintains his indifference to the lower interest rate and increased (relative to pre-intervention) loan size of the LR.

There is an alternative way of looking at our result. We find that a subsidy can be Pareto improving because the envelope theorem does not hold for screened borrowers - a change in their loan rate has both direct and indirect effects on their indirect objective function in the competitive equilibrium (CE). This should not be surprising because their indirect profit function does not represent the *maximum* level of profits but the constrained maximum. Although loan qualities are not bundled at a single price as would occur in a pooling equilibrium, the screening CE has only determined quality-specific prices by the creation of an externality - the loans received by screened borrowers reflect the characteristics of others in the market and the marginal rates of substitution (MRS) between borrower and bank are not equated. If we regard the loans received by different types as different factors of production, we see that this implies that the MRS between different inputs are not equated to relative price and there are therefore loan reallocations, achieved via tax-induced changes in relative rates, that can be Pareto-improving. Given no more information than the banks a central planner can improve on the decentralized system which, constrained not to make a loss, lacks the planner's redistributive ability.

As discussed in Chapter 7.3 and Chapter 2, McKinnon and Shaw originally[1] argued that allowing banks to determine loan rates free of regulation would permit them to exercise their comparative advantage in information acquisition, allocating loans according to expected social return. We argued in Chapter 7 that screening supports MS's intuition that financial intermediaries have arrangements which help alleviate the inefficiencies caused by imperfect information. This chapter demonstrates that they were over-optimistic in their assessment of outcomes in a deregulated market. Even sophisticated banking arrangements may be improved upon by careful policy.

Before deriving this result we discuss the policy measures already suggested in the literature. Section 8.2 considers the policy issues raised when banks use only the loan rate. It

---

[1] We will discuss McKinnon's [1988] revision of this view.

compares the results of models with similar structures but differing in their bank contract assumptions to conclude that policy measures should take account of existing intermediary arrangements. Brief consideration is also given to the effects stabilization policy may have in credit markets with adverse selection. Section 8.3 examines welfare issues in markets where more sophisticated banks are acquiring information. In the main it discusses the Greenwald-Stiglitz analysis and our results. Only de Meza and Webb's [1988] paper examines policy when banks actively acquire information and there the information is acquired through a credit rating agency. In Section 8.4 the partial equilibrium model of Chapter 7 is embedded in the adapted Greenwald-Stiglitz general equilibrium framework to show that Pareto-improving official action is consistent with the most rational and undistorted behaviour postulated for information-constrained banks.

8.2: Policy in models with loan rate contracts

Where models of asymmetric information in the credit market consider or derive contracts specifying only the loan rate, various policy measures have been suggested to deal with the resulting distortions in investment finance. The social loss in this case results from the fact that, while society's interest requires maximization of the expected return from project implementation, bank strategy, concerned only with the return in the event of success, may either exclude high expected return borrowers (rationing) or permit a loan rate which attracts projects with expected return different from the opportunity cost of capital.

In the context of the Stiglitz-Weiss [1981] rationing of observationally distinct groups, Ordover and Weiss [1981] suggest that banks be forced to lend to all borrowers so that the inclusion of projects with high but risky returns raise the expected economy-wide return per dollar lent[2]. Such total removal of discretion from banks is impractical because it would interfere with any evaluation/assessment techniques the banks have evolved, as well as nullifying all contract possibilities. Further, borrowers' knowledge of such regulation seems likely to create its own adverse

---

[2] The supply of loanable funds would be maintained unchanged through proportional taxes on borrowers' profits and transfers to depositors.

selection effects: for borrowers with poor repayment prospects and protected by limited liability such a rule would present a golden opportunity.

de Meza and Webb's [1987] model demonstrates the inefficiencies caused by adverse selection. Nevertheless, a comparison of models indicates that MS's warning about government intervention should be noted; in some cases it is not clear that the policy measure proposed would be necessary if allowance is made for sorting possibilities. de Meza and Webb model a competitive banking industry whose potential borrowers are ranked by first order stochastic dominance (FOSD). Although there is no rationing since bank return is monotonic increasing in the loan rate, competition reduces the loan rate to the point where banks finance projects whose expected return is below the opportunity cost of loanable funds: there is overinvestment[3]. First-best is achieved by a tax on depositors' interest income. This raises the deposit rate banks must pay and hence increases the loan rate at which the zero-profit market-clearing equilibrium occurs, deterring those HR borrowers with negative expected social returns. Comparison of this model with Besanko and Thakor's [1987a] competitive model (see Chapter 7.2) suggests that where banks compete in contracts which specify collateral as a screening device, the market alone can eliminate such overinvestment. Collateral provides banks with a means of distinguishing risky borrowers and charging the HR a higher rate of interest.

When de Meza and Webb use the original Stiglitz-Weiss risk definition, they not only find that there may be a credit rationing equilibrium, but that even in its absence there may be underinvestment. At higher interest rates the entry of HR projects implies that banks must charge a sufficiently higher rate of interest to compensate for the increased risk. Consequently, competitive equilibrium occurs where expected social project return exceeds the opportunity cost of project investment (the deposit rate). The rate charged by banks must compensate them for risk at a rate that exceeds the social valuation because they are concerned only with success return[4]. In

---

[3] Recall that with FOSD the adverse selection effect operates so as to attract HR (low expected return) borrowers as the loan rate decreases.

[4] The underinvestment result is less acute when credit is rationed, illustrating the social benefits of even this bank response. With rationing the negative adverse selection effect on expected bank profit which occurs at higher interest rates is reduced so banks' interest rates are nearer to expected project return.

this case a subsidy on interest income, by reducing the deposit rate banks must pay, allows them to expand loans to the level necessary for first best investment. If, however, the sorting possibilities available to banks are considered adequate, comparison of this model with that of Bester [1985] would suggest that the adverse selection effect responsible for underinvestment could again be overcome by screening contracts with collateral.

Since collateral is costly when borrowers are risk averse, Greenwald and Stiglitz's analysis and the results of our application in Section 8.4 suggests that a planner could improve on the screening equilibrium. Furthermore, the taxes and subsidies suggested by de Meza and Webb seem likely to be the appropriate instruments in the risk environments they specified. However, an investigation of policy with collateral screening would be needed for a clear indication.

The juxtaposition of these underinvestment and overinvestment results stresses the commodity-bundling effects of imperfect information. The basic problem is that qualities cannot be identified and correct prices determined. Heterogeneity and private information causes social and private (bank) returns to diverge whatever the risk definition used, the definition determining only whether the selection is adverse or "proverse" to the bank. In a pooling equilibrium HR and LR borrowers pay the same interest rate which reflects only the average risk on the market. If expected return and success probability are positively correlated (FOSD) this results in too many projects being financed[5], if non-positively (SOSD) too few projects are financed.

Other implications of adverse selection of special significance to LDC monetary policy have been discussed. McKinnon [1988] has recently recognized the possibility of equilibrium credit rationing due to asymmetric information. As a result of this recognition and the further assumption of macroeconomic instability, he virtually reverses his earlier recommendation of minimal regulation. He suggests that macroeconomic instability can produce positive yield correlation within groups so that whole borrower classes may be forced into default in adverse macro situations, increasing the bankruptcy risk of the lending bank. In these conditions

---

5 Comparison of de Meza and Webb's analysis, which uses the zero profit result of a competitive equilibrium to show underinvestment, with that of Stiglitz and Weiss's suggests that the rationing result is more likely if banks are assumed to behave strategically, recognizing the adverse effects of interest-rate increases past the bank-optimal rate. Myopic competition seems most likely to result in higher rates.

McKinnon suggests that it may be appropriate for the authorities to impose an interest rate ceiling. This would both discourage banks from lending to the riskiest borrowers and reduce a borrower's incentive to choose riskier, higher return investments. But interest rate regulation does not appear well-justified here. In the circumstances described it is not clear why the authorities should be more adept at (concerned with) preventing its bankruptcy than the bank itself. A profit-maximizing bank, aware of the positive covariance among returns in a particular risk class, would wish to structure its overall loan portfolio so as to minimize bankruptcy costs, if its bears the entirety of those costs. On the other hand, and this is the major focus of McKinnon's new analysis, banks' belief in a government commitment to rescue them in the event of bankruptcy does pose a moral hazard problem. With a rescue system in place, banks have an incentive to lend to riskier borrowers, at higher interest rates than they otherwise would, because they no longer bear the full cost of loan default. It is in this context that the new McKinnon advocacy of an interest rate ceiling regulation may be justified. Assuming that the authorities have the same information as the banks, the regulated rate should be set at the bank-optimal rate in the absence of moral hazard. However, in this change of direction McKinnon does not take account of banks' screening and incentive arrangements. If discriminating contracts are offered by banks, a blanket interest rate ceiling of the type he suggests, may interfere with their ability to sort and result in the exclusion of investments that may be socially desirable ... as McKinnon originally argued. This is easily seen by considering the result of a ceiling below R_2 in the model of Chapter 7.

McKinnon placed his analysis in the context of the macroeconomy, implicitly recognizing that when individual behaviour does not satisfy once-standard assumptions, its interaction with macroeconomic events may give surprising results. We have already described in Section 7.2 Blinder's [1987] analysis of the enhanced effects of expansionary monetary policy. Mankiw [1986] has considered the converse case of restrictive monetary policy. He considers an exogenous upward shift of banks' supply curve which raises the loan rate. Adverse selection may then decrease the average success probability faced by banks, necessitating a further increase in loan rate. This process could lead, in Mankiw's words, to 'financial collapse' - the pool of borrowers available to banks becomes so risky that no loans are made. A major plank of

stabilization policies is restrictive monetary policy aimed at reducing demand. It assumes that only marginally productive investment projects will be precluded, and that the funds released by reduced government borrowing and attracted by higher interest rates will permit financing of the most productive investments, which can pay a higher rate of interest. But no account is taken of the divergence between social and private returns induced by private information. The adverse selection effect could result in investment finance stagnation, despite the availability of projects which are socially productive at the higher rates of interest. While this does not seem a good description of current conditions in industrialized countries - Mankiw mentions historical incidents of instability - it does not seem too far removed from conditions in several LDCs. Mankiw suggests that restrictive policy be accompanied by special government assistance, such as guarantees[6], to certain borrowers. Stabilization programmes may have to include quite active government intervention, in the form of official project evaluation, subsidies and loan guarantee schemes, in order to obtain socially productive investment. Mankiw does not consider sorting contracts, however, and their absence or presence would need to investigated.

8.3: Policy when Banks are Active Information Processors

We first consider the only published model which considers tax policy when banks actively obtain information, before discussing the general approach used by Greenwald and Stiglitz which we shall apply in the following section.

de Meza and Webb [1988] examine the role of tax policy where banks obtain information directly rather than through self-selection contracts: banks pay a fixed fee to a credit rating agency to learn the expected return on individual projects. Clients choose whether to be rated. Clearly this choice is itself a screen and rating is chosen by the LR. But there is still overinvestment because the adverse selection effect continues to operate in the unrated group where the interest rate reflects average success probability. But now there is the additional social cost of the rating. A prohibitive tax on screening and a tax on depositor income would give the first best

---

[6] Mankiw notes that a subsidy serves a function similar to a loan guarantee which removes the downside risk for banks.

result.

Greenwald and Stiglitz argue that where information is asymmetric markets are not constrained Pareto efficient because the quality effects of price have results similar to technological externalities. The overinvestment, underinvestment[7] and collateral inefficiencies which result in the screening models discussed in Chapter 7.2 may be seen as representative of this externality. The source of the externality described by Greenwald and Stiglitz appears to lie in the fact that the average riskiness on the market affects the individual borrower (except for the marginal) who must accept a contract which reflects not his characteristics but market-imposed constraints.

Greenwald-Stiglitz (GS) make two points. In the first place they remind us that pecuniary externalities have welfare consequences when the economy is distorted. And, as the results discussed above illustrate, asymmetric information does produce distortions. Secondly, that the distortions produced by asymmetric information or incomplete markets resemble technological externalities. GS offer a general methodology which exploits this similarity, introducing a quality variable into the utility and production functions of a general equilibrium model. The criterion used to determine whether a Pareto improvement is possible in the presence of asymmetric information (with or without sorting/incentive-compatible contracts) is whether, at the competitive equilibrium (CE) with no intervention, there exist taxes or subsidies and lump-sum transfers that can improve everyone's welfare.

Operationally, in an assumed intervention-free CE they investigate whether there is a set of taxes etc. such that household utilities are unchanged and government net revenues increased. If the change in government net revenue is zero when taxes are zero, having taken account of how a tax change affects equilibrium decisions via the indirect value functions, the original equilibrium is constrained Pareto efficient. That is, any government intervention which leaves household utilities unchanged has no effect on government revenue, and hence a Pareto improvement is not possible with taxes/subsidies and lump-sum transfers alone. In general, however, taxes change the level of the unobservable variable chosen by market participants (moral hazard case), or the

---

[7] These are relative to symmetric information, result from screening by loan size and should not be confused with the inefficient investment from adverse selection.

average quality of those on the market (adverse selection), or the level of signal chosen (screening/signalling). The tax change, which changes prices, is the source of the pecuniary effect, its allocative and distributive effects are zero at the CE, but the change in the level of the unobservable variable (like the technological externality) affects government revenue.

Our results differ in flavour from G-S's. As indicated above, they find a Pareto improvement because the tax-induced change in their quality variable is not eliminated by the envelope theorem. We find that the envelope theorem does not apply to screened borrowers because the incentive compatibility constraint forces them to be away from their profit maximizing loan level. The scope for government intervention appears to arise from the fact that MRS are not equated. The G-S externality effect is operating via the fact that profit functions at the CE are not separated (except for the highest risk who is not signalling), since it is only because types are linked through the incentive constraints that identification is possible.

Intuition suggests that it may be possible to consider the G-S analysis in the light of recent research on the Pareto efficiency of missing markets. This possibility is suggested by our finding that non-equalization of MRS[8] is the source of a Pareto improvement. Similarly, the non-equalization of MRS is the intuitive explanation for Geanokoplos and Polemarchakis's [1986] finding of generic suboptimality of equilibria with incomplete markets. They proved that when asset markets are incomplete a CE exists but that the equilibrium portfolio allocation is almost everywhere constrained inefficient - the structure of markets in the CE is not used efficiently. When markets are complete changes in relative prices have no effect on actions and welfare because the envelope theorem applies. If there are markets missing, however, asset reallocations which change relative prices affect welfare generically. Since a technological externality essentially results from a missing market, G-S's analogy should be consistent with the more formal analysis.

---

[8] We may note that Gale [1987] overcame the problem of missing markets to obtain a Walrasian equilibrium in his model of adverse selection in the credit market by the devices of complete contracts and a matching game - the latter giving MRS equality.

8.4: The General Equilibrium Framework

In this section, the partial equilibrium model developed in Chapter 7 is embedded in G-S's general equilibrium framework in order to investigate whether a tax/subsidy and lump-sum transfer scheme can bring about a Pareto improvement, at least theoretically.

The usual assumption for economies with asymmetric information is that, with no better information than the market, the authorities can do no better. The point of the G-S methodology would appear to be its focus on government intervention via taxes/subsidies (together with lump-sum transfers) which require only that the government have the same information as banks. The tax-induced price changes, by changing agents' opportunity costs, change the quality choice and hence bring about a Pareto improvement. However, although G-S sketch several models to illustrate applications of their general approach, their results are not readily interpretable in terms of the usual partial equilibrium models of markets with asymmetric information. Most extant models of screening/signalling etc. incorporate indivisible variables/ uncertainty conditions that do not fit into the neoclassical general equilibrium framework. Hence the use of the partial equilibrium model developed in Chapter 7. Although our use of it has been slightly complicated, at least notationally, by the wish to take account of the existence of observationally distinct groups which, Section 7.3 suggests, is usually taken to be a feature of LDC markets.

The fact that our concern is with a credit market complicates the use of the G-S framework. We need to introduce a credit market and banking sector into the general equilibrium framework. However, the microeconomic foundations of a monetary economy are not well -developed. A logically specified model consistent with the existence of money and financial institutions would require the explicit introduction of time, state contingency, strategic behaviour and issues of trust; see, for example, Gale [1982] and Shubik [1987]. In order to avoid the modelling complications accompanying these factors, the following assumptions are made:

Assumptions

(1) Banks obtain their loanable funds from the household endowment so that "deposits" are viewed as an endowment good which may be either consumed by the household or invested with

banks for a return - sold to banks for a price which is the deposit rate factor. While this formulation captures the notion of a tradeoff between the consumption and deposit decisions, it does not reflect the intertemporal nature of the choice.

(2) The assumption of marginal cost pricing by banks was discussed in Chapter 7.

(3) Loans supplied by banks are treated as a variable input in firms' production functions.

(4) The justification for banks in our framework is their assumed ability to offer screening contracts; an ability which the household who owns the endowment does not possess. A drawback is that this ability is exogenously imposed, rather than endogenously determined, however the aim is not to explain the existence of banks.

(5) It is the combination of uncertainty regarding project yield and the asymmetry of information regarding project type that leads banks to use sorting contracts. In order to avoid aggregate uncertainty, it is assumed that, although information regarding the identities of individual agents (firms) in the screened sector is private, there is public knowledge of output realizations.

The general equilibrium framework used follows that of Greenwald and Stiglitz very closely. It differs in the following respects: a) A banking sector is introduced. b) A single representative household is assumed and given an endowment which serves as the deposit good. The assumption of a representative household saves on notation in our model. Endowments are excluded in their general model as a simplification. c) Individual firm output is random. d) Production sectors are distinguished.

We proceed by a) describing the economy and assuming a competitive equilibrium with no government intervention; the important feature of this description is the specification of indirect objective functions and their price derivatives at the CE, and b) testing for a tax and transfer policy which changes government revenue at a zero tax rate while leaving profits and utility unchanged. The economy's agents consist of firms, households, banks and government. The following describes their maximization problems and the outcomes.

FIRMS

Firms are distinguished by their management 'type', θ_f. Each firm maximizes expected profit

$$E\phi_f = y_{1f} + q_f \sum_{j=2}^{N} p_j y_{jf} \tag{8.1}$$

subject to the production constraint:

$$y_{1f} - G_f(q_f, Y_f, \theta_f) \leq 0$$

where

ϕ_f = profit function for firms of type θ_f,

(y_{1f}, Y_f) = production vector for firm of type f,

y_{1f} = numeraire production good,

$Y_f = (-L_f, y_{3f}, \ldots, y_{Nf})$ is a vector of the N-1 non-numeraire production goods,

   $y_{jf} < 0$ denotes that good j is an input to a firm of type f,

L_f = loan input to firm of type θ_f,

$P = (R, p_3, \ldots, p_N)$ = vector of producer prices for $N-1$ non-numeraire goods,

R = loan rate factor in terms of the numeraire good,

q_f = success probability of firm of type θ_f,

$G(.)$ = production function,

θ_f = firm-specific factor such as management ability.

γ_f = proportion of firms of type θ_f

γ_{fi} = proportion of firms of type θ_f producing one or more of the goods $i = 3, \ldots, k-1, k+1, \ldots, N$

y_k = output of only sector subject to asymmetric information

In the production vector, good k, y_{kf}, with production conditions specified in the partial equilibrium model of Chapter 7.4, is distinguished. In line with the idea that there are observationally distinguishable groups in the credit market, y_{kf} is interpreted as the output of a non-traditional sector or industry. In sector k each firm's θ_f, and hence q_f is private knowledge.

It may be assumed that all industries use L as an input. In other sectors (industries) banks have *a priori* knowledge of each firm's θ_f and q_f i.e. they can identify the firms individually. Therefore screening only takes place in the y_k sector. Other firms and their outputs are indexed i. The production functions of non-y_k firms need not be specified but, unlike the particular production and profit functions used in the partial equilibrium model, firms not producing y_k may be assumed to have factor substitution possibilities. For these sectors it is assumed that y_{kf} is an input if $y_{kf} \neq 0$.

We assume that the distribution of firm types and the production functions are public knowledge. For notational simplicity we maintain Section 7.4's assumption that there are only two types in the economy, but that there are many firms of each type. In order to avoid problems of aggregate uncertainty, we make the assumption commonly used for this purpose: that it is public knowledge that a proportion $(1 - q_f)$ of the fraction γ_f of each type θ_f will fail, and γ_f is known. We use the normalisation that output is zero in the event of failure. Thus total production of any good j is:

$$y_j = q_1 \sum_i \gamma_{1i} \, y_{j1i} + q_2 \sum_i \gamma_{2i} \, y_{j2i} \tag{8.3}$$

where y_{j1i} is the production of good j by a firm of type θ_1 producing outside of the k industry, and $q_1 \gamma_{1i}$ is the share of those firms of type θ_1 producing good y_j whose production process is successful. Aggregate production is thus known and in a perfectly competitive system, prices may be treated as if parametric to all agents. (While R is an exception, as discussed earlier the Bertrand assumption implies that the loan rate is determined at the marginal cost level.) Private information pertains only to individual firm identification by type in industry k. This assumption to avoid the effects of uncertainty is formally correct only in a continuum economy where the influence of a finite number of agent types can be smoothed by aggregation. In order to retain the flavour of this approach use has been made of firm shares. Then γ_f should be interpreted as the share of all agents who are firms of type θ_f. Similarly, since banks are all assumed to have the same profit functions, the share of agents acting as banks could be denoted γ_b.

Profit and demand functions of i firms

Since all firms receive loans, with banks being compensated for risk by the interest rate charged, total loans to firms of type θ_f in the sectors i are:

$$q_f \sum_i \gamma_{fi} \, L_{fi} + (1-q_f) \sum_i \gamma_{fi} \, L_{fi} = \gamma_{fi} \, L_{fi} \tag{8.4}$$

where L_{fi} = loan size to firm of type θ_f producing i. For firms in the i sectors the realized indirect profit function is given by

$$\overset{*}{\phi}_{fi} (R_f, p_k, p^*)$$

where R_f = interest rate charged to firm of type θ_f; p_k = price of good y_k, if y_k enters these firms'

profit functions it is assumed to do so as an input; p^n = vector of prices of goods other than y_k. As usual,

$$\frac{\partial \overset{*}{\phi}_{fi}}{\partial p_n} = y_{nfi} \tag{8.5}$$

is the commodity supply of or (if negative) factor demand for good n by a firm of type θ_f in an i industry.

Profit and demand functions of k firms

From the partial equilibrium model, the following indirect profit function is assumed to characterize firms in the y_k sector:

$$\overset{*}{\phi}_{fk} = [p_k (L_1^c)^\beta V(\theta_f) - R_f L_f^c] \tag{8.6}$$

We know from (7.14iibis) that for firms of type θ_2 who receive their first-best contract the loan received is

$$L_2^c = \left[\frac{p_k \beta V(\theta_2)}{R_2} \right]^{\frac{1}{1-\beta}}$$

where L_f^c = screening CE loan size to firm of type θ_f. Differentiating (8.6)

$$\frac{\partial \overset{*}{\phi}_{2k}}{\partial R_2} = \left[p_k \beta \frac{\partial L_2^c}{\partial R_2} (L_2^c)^{\beta-1} V(\theta_2) - L_2^c - R_2 \frac{\partial L_2^c}{\partial R_2} \right]$$

Substituting for $\frac{\partial L_2^c}{\partial R_2} = -\frac{L_2^c}{R_2} \frac{1}{1-\beta}$ and simplifying:

$$\frac{\partial \overset{*}{\phi}_{2k}}{\partial R_2} = -L_2^c \tag{8.7}$$

which is the demand for the loan factor by firms of type θ_2. Similarly,

$$\frac{\partial \overset{*}{\phi}_{2k}}{\partial p_k} = y_{k2k} \tag{8.8}$$

is the supply of y_k by firms of type θ_2. Note that the envelope theorem holds as expected for these types who do not signal.

For the firms of type θ_1 in the k industry who are signalling their low risk L_1^c is defined by ((see 7.15ii) and note (7.24)) :

$$p_k (L_2^c)^\beta V(\theta_2) - R_2 L_2^c - p_k (L_1^c)^\beta V(\theta_2) + R_1 L_1^c = 0 \tag{8.9}$$

Thus, unlike firms of type θ_2, type θ_1s' loan size and indirect profit function are dependent on type θ_2's contract and management parameter. Differentiating (8.6) for type θ_1 with respect to input and output prices, we find that, since θ_{1k} is not producing according to its own-optimum, not only does a change in price have a larger than direct effect on its profit function, but that its profit function is affected by changes in the loan rate charged to θ_2:

$$\frac{\partial \overset{\bullet}{\phi}_{1k}}{\partial R_1} = \left[p_k \beta (L_1^c)^{\beta-1} V(\theta_1) - R_1 \right] \frac{\partial L_1^c}{\partial R_1} - L_1^c \qquad (8.10)$$

where, implicitly differentiating (8.9),

$$\frac{\partial L_1^c}{\partial R_1} = \frac{L_1^c}{\beta p_k (L_1^c)^{\beta-1} V(\theta_2) - R_1} > 0 \qquad (8.11)$$

The inequality in (8.11) follows from (7.23): $\beta p_k (L_1^c)^{\beta-1} V(\theta_1) - R_1 > 0$, and $V(\theta_2) > V(\theta_1)$ so that the denominator in (8.11) is positive. Thus an increase in R_1 has the usual direct effect of negative factor demand on the indirect profit function, plus a positive effect whose size depends on the differences between the marginal productivity value of a loan and its marginal cost which arise because of the management parameter. Differentiating (8.6) for θ_1 with respect to R_2:

$$\frac{\partial \overset{\bullet}{\phi}_{1k}}{\partial R_2} = \left[p_k \beta (L_1^c)^{\beta-1} V(\theta_1) - R_1 \right] \frac{\partial L_1^c}{\partial R_2} \qquad (8.12)$$

where

$$\frac{\partial L_1^c}{\partial R_2} = - \frac{L_2^c}{\beta p_k (L_1^c)^{\beta-1} V(\theta_2) - R_1} < 0 \qquad (8.13)$$

Since an increase in R_2 reduces the loan size demanded by type θ_2, it implies that the loan size required to distinguish type θ_1 must fall, with a negative effect on the latter's profit - θ_1 must move to a higher isoprofit curve in $<R,L>$ space. Considering the effect of a marginal change in p_k on θ_1's equilibrium profit:

$$\frac{\partial \overset{\bullet}{\phi}_{1k}}{\partial p_k} = y_{k1k} + \left[p_k \beta (L_1^c)^{\beta-1} V(\theta_1) - R_1 \right] \frac{\partial L_1^c}{\partial p_k} \qquad (8.14)$$

where

$$\frac{\partial L_1^c}{\partial p_k} = \frac{V(\theta_2) [(L_2^c)^\beta - (L_1^c)^\beta]}{\beta p_k (L_1^c)^{\beta-1} V(\theta_2) - R_1} > 0 \qquad (8.15)$$

Thus an increase in p_k has the usual direct effect of the supply of the produced good, plus a

positive indirect effect which appears to reflect the increase in θ_2's output and hence loan demand (allowing θ_1 to accept a larger loan) following an increase in p_k.

The externality resulting from asymmetric information is evident. Compare the effects on θ_2 and θ_1 types of a change in prices. For θ_2 who receive the first-best contract, loan size depends only on their own quality and is identical to loan size demanded. But, as a result of the incentive constraints, there is a type of jointness in production for type θ_1s and the envelope theorem does not hold. In equations (8.10), (8.12) and (8.14) the marginal effect of price differs from the usual direct effect because $dR(\hat{\phi}_{1k})/dL > 0$ in the screening equilibrium - see equation (7.23) - which results from the binding incentive compatibility constraint on type θ_2 in industry k. Note that in the many-type economy, with many probabilities of success, only the final type with lowest probability would be receiving the first-best contract.

REPRESENTATIVE CONSUMER

The representative household maximizes the utility function:

$$U = U(x_1, x_d, X),$$

subject to:

$$x_1 + \sum_{n=3}^{N} c_n x_n \leq T + \sum_{f=1,2} q_f \gamma_{fk} \phi_{fk} + \sum_r \sum_{f=1,2} q_f \gamma_{fi} \phi_{fi} + I(\bar{x}_d - x_d) \tag{8.16}$$

where:
x_1 = consumption of numeraire good
x_d = consumption of deposit good
$\hat{x}_d$ = compensated demand for x_d
$\bar{x}_d$ = household's endowment of the deposit good
x'_d = net supply of deposits
X = consumption vector of $N-2$ non-numeraire and non-deposit goods
x_n = consumption of good n
c_n = consumer price of good n in terms of numeraire good
c^n = vector of consumer prices for goods other than k
c_k = consumer price for good x_k in terms of numeraire good
T = lump-sum government transfer to household, when operable
$q_f \gamma_{fk} \phi_{fk}$ = total realized profits for firms of type θ_f in the k sector
I = deposit rate factor in terms of numeraire good

The solution to the household maximization problem gives the expenditure function:

$$E(I, c^n, c_k, u^*)$$

where $u^* = $ the utility level attained in the competitive equilibrium with screening and without government intervention. Differentiation of the expenditure function gives the compensated demand for a commodity, thus

$$\frac{\partial E}{\partial I} \mid u^* = \hat{x}_d (I, c^n, c_k, u^*) \tag{8.17}$$

is the compensated demand function for x_d as a consumption good. The household can either consume its endowment, $\bar{x}_d$, or sell to banks in return for payment I. The uncompensated demand is $x_d (I, c^n, c_k, \bar{x}_d, T)$, ($= \hat{x}_d$ in CE), so the net supply of deposits to the banking system is:

$$[\bar{x}_d - x_d (I, c^n, c_k, \bar{x}_d, T)] = x'_d (I, c^n, c_k, \bar{x}_d, T) \tag{8.18}$$

BANKS

Banks function only as conduits of deposit endowments from household to firms and as contract writers, where contract writing is assumed a costless activity. Banks with separable profit functions maximize:

$$\sum_F E \pi_f = q_f R_f L_f - I L_f \tag{8.19}$$

where $\sum_F$ indicates summation over all types in all sectors. It is assumed that all banks offer the same contracts and that borrowers and depositors are randomly allocated among the banks. Since the model in Chapter 7 shows that interest-rate setting in contracts by banks results in marginal cost pricing, bank profits are zero and do not affect the household budget constraint.
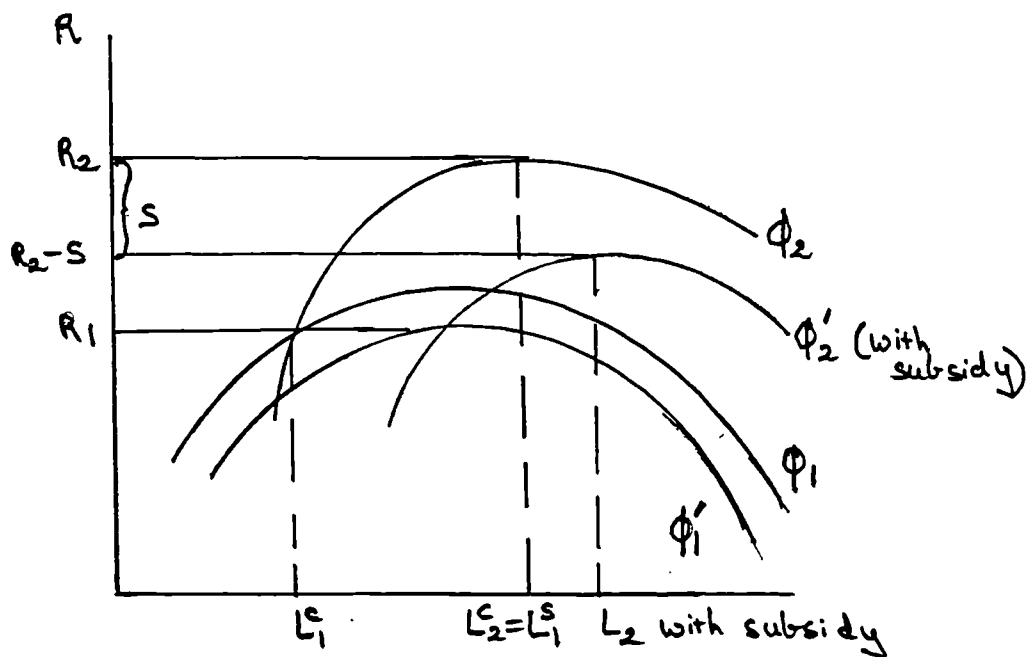
GOVERNMENT

The government collects taxes from (or gives subsidies to) firms borrowing from banks and makes lump-sum transfers to (imposes lump-sum taxes on) the household who owns the firms, receiving net income:

$$v = t q_2 \left[\gamma_{2k} L_{2k} + \sum_i \gamma_{2i} L_{2i} \right] - T \tag{8.20}$$

where $t = $ the tax imposed on loans with interest rate $R_2 = I / q_2$. Negative t implies that loans to this sector are subsidised, T is then also negative, a lump-sum tax imposed on the household.

This formulation of the government's maximand is based on the partial equilibrium model. It is type θ_2's loan size which constrains the loan size obtained by the types θ_1. The smaller is the loan size taken by type θ_2, the smaller the loan size required for θ_1 to signal higher success probability. Intuitive reasoning then suggests that reducing the loan size taken by types θ_2 would increase the θ_1 loan size banks can offer while continuing to induce self-selection. This is illustrated in Figure 8.1. A subsidy of s on the higher loan rate moves θ_2 to a lower isoprofit (where profit is higher) curve where he is just indifferent between the subsidized contract he receives and the contract chosen by θ_1. The latter is θ_1's first-best contract. Although subsidizing θ_2 has shifted him from his original first-best position, he remains on his demand curve.

Figure 8.1: *A Pareto-improving loan rate subsidy*

What advantage does the government have here that banks do not have? Banks are as aware of this possibility as are the authorities but, given the market deposit rate of interest, and the success probabilities of the two types, banks cannot charge θ_2 a rate below R_2 without making a loss. A bank could only afford to do so if it were able to 'tax' some other group receiving credit. But raising the loan rate for that group would induce its departure to other banks with a lower loan rate. Essentially, the authorities have 'market power' that, in the perfectly competitive system postulated, banks do not.

Government intervention in the assumed competitive equilibrium

To test our intuition using the Greenwald-Stiglitz criterion, it is assumed that a competitive equilibrium, without government intervention, exists, i.e. $T = 0 = t$. In that competitive equilibrium producer prices equal consumer prices:

$$(R_f, P^-) = (q_f^{-1} I, C)$$

where: $P^- = (p_3, \ldots, p_N)$ = producer price vector, omitting the price of the loan factor,

$C = (c_3, \ldots, c_N)$ = consumer price vector for non-deposit goods

and all markets clear: the demand for all goods n, $n = k, i$ ($i = 3, \ldots, k-1, k+1, \ldots, N$), equals the realized supply of good n over all firm types:

$$\sum_n x_n(I, c, \bar{x}_d) = \sum_f \sum_j \gamma_{fj} \, q_f \, y_{nfj}(R, P^-, \theta_f) \qquad (8.21)$$

and the net supply of the deposit good by the household equals the total loans supplied:

$$x'_d(I, c, \bar{x}_d) = \sum_f \sum_j \gamma_{fj} \, L_{fj}(R, P^-, \theta_f) \qquad (8.22)$$

The government maximizes revenue subject to the constraint that household utility remain at the level in the CE with screening and without government intervention. It therefore solves the following problem:

$$\underset{t,T}{Max}\ v\ =\ tq_2[\gamma_{2k}L_{2k} + \sum_i \gamma_{2i}L_{2i}] - T \tag{8.23}$$

subject to $E[I(t),c^*(t),c_k(t),u^*] = T(t) + I(t)\bar{x}_d$

$$+ q_1\gamma_{1k}\phi_{1k}\Big[R_1(t),R_2(t)+t\ p_k(t),\theta_1\Big]$$

$$+ q_2\gamma_{2k}\phi_{2k}\Big[R_2(t)+t\ p_k(t),\theta_2\Big]$$

$$+ q_1\sum_i\gamma_{1i}\phi_{1i}\Big[R_1(t),p_k(t),p^*(t),\theta_1\Big]$$

$$+ q_2\sum_i\gamma_{2i}\phi_{2i}\Big[R_2(t)+t\ p_k(t),\theta_2\Big]$$

That is, government maximizes its net revenue subject to maintaining the expenditure of the consumer at the competitive equilibrium level given by its income from the transfer, if any, income from the deposit good and from firm profits. If the original CE is Pareto efficient, this problem has a zero solution at $t = 0$. That is, taxes (subsidies) do not exist such that government revenue can be marginally increased while maintaining household utilities unchanged.

If the constraint in (8.23) is substituted into the maximand, a maximization problem can be solved in its usual form. Notation is simpler, however, if the constraint is first considered separately. Along the constraint, the change in lump sum transfer, for a unit change in t, required to keep the household at utility level u^* is:

$$\frac{dT}{dt} = \frac{\partial E}{\partial I}\frac{dI}{dt} + \frac{\partial E}{\partial c_k}\frac{dc_k}{dt} + \sum_n\frac{\partial E}{\partial c_n}\frac{dc_n}{dt} - \bar{x}_d\frac{dI}{dt} \tag{8.24}$$

$$- q_1\gamma_{1k}\left[\frac{\partial\phi_{1k}}{\partial R_1}\frac{dR_1}{dt} + \frac{\partial\phi_{1k}}{\partial R_2}\frac{dR_2}{dt} + \frac{\partial\phi_{1k}}{\partial p_k}\frac{dp_k}{dt}\right]$$

$$- q_2\gamma_{2k}\left[\frac{\partial\phi_{2k}}{\partial R_2}\frac{dR_2}{dt} + \frac{\partial\phi_{1k}}{\partial p_k}\frac{dp_k}{dt}\right]$$

$$- q_1\sum_i\gamma_{1i}\left[\frac{\partial\phi_{1i}}{\partial R_1}\frac{dR_1}{dt} + \frac{\partial\phi_{1i}}{\partial p_k}\frac{dp_k}{dt} + \sum_n\frac{\partial\phi_{1i}}{\partial p_k}\frac{dp_k}{dt}\right]$$

$$- q_2\sum_i\gamma_{2i}\left[\frac{\partial\phi_{2i}}{\partial R_2}\frac{dR_2}{dt} + \frac{\partial\phi_{2i}}{\partial p_k}\frac{dp_k}{dt} + \sum_n\frac{\partial\phi_{2i}}{\partial p_k}\frac{dp_k}{dt}\right]$$

Using the envelope theorem i.e. equations (8.5), (8.7) and (8.8), together with equations (8.10) to (8.15), and noting $\frac{d\,p_k}{d\,t} = \frac{d\,c_k}{d\,t}$ at zero taxes and $\frac{dR_2}{dt} = \frac{1}{q_2}\frac{dI}{dt} + 1$ (8.24) can be written:

$$\frac{dT}{dt} = -(\bar{x}_d - x_d)\frac{dI}{dt} \tag{8.25}$$

$$+ \frac{dp_k}{dt}\Big[x_k - q_1\gamma_{1k}y_{k1k} - q_2\gamma_{2k}y_{k2k} + q_1\sum_i\gamma_{1i}y_{k1i} + q_2\sum_i\gamma_{2i}y_{k2i}\Big]$$

$$+ \sum_n\frac{dp_n}{dt}\Big[x_n - q_1\sum_i\gamma_{1i}y_{n1i} - q_2\sum_i\gamma_{2i}y_{n2i}\Big]$$

$$+ \frac{dI}{dt} \left[\gamma_{1k} L_{1k}^c + \gamma_{2k} L_{2k}^c + \sum_i \gamma_{1i} L_{i1} + \sum_i \gamma_{2i} L_{2i} \right]$$

$$+ q_2 \gamma_{2k} L_{2k}^c + q_2 \sum_i \gamma_{2i} L_{2i}$$

$$- q_1 \gamma_{1k} A \left\{ \frac{\partial L_{1k}^c}{\partial R_1} \frac{dR_1}{dt} + \frac{\partial L_{1k}^c}{\partial R_2} \frac{dR_2}{dt} + \frac{\partial L_{1k}^c}{\partial p_k} \frac{dp_k}{dt} \right\}$$

where $A = p_k \beta (L_1^c)^{\beta-1} V(\theta_1) - R_1$, that is, the difference between the value of marginal productivity of the loan to types θ_1 in sector k and the price of the loan. Note that y_{k1k} is the supply of good k by firms of type 1 in the k industry, y_{k1i} is the demand for good k by firms of type 1 in the i industry, and y_{n1i} is the supply of good n by firms of type 1 in the i industry. Market clearing, (8.21), implies that the square bracketed terms above are zero, and since $(\bar{x}_d - \hat{x}_d) = x'_d$ is deposit supply, with clearing in the loan market equation (8.25) becomes:

$$\frac{dT}{dt} = q_2 [\gamma_{2k} L_{2k}^c + \sum_i \gamma_{2i} L_{2i}] \tag{8.26}$$

$$- q_1 \gamma_{1k} A \left\{ \frac{\partial L_{1k}^c}{\partial R_1} \frac{dR_1}{dt} + \frac{\partial L_{1k}^c}{\partial R_2} \frac{dR_2}{dt} + \frac{\partial L_{1k}^c}{\partial p_k} \frac{dp_k}{dt} \right\}$$

Differentiating the government maximand and substituting from equation (8.26) for $\frac{dT}{dt}$

$$\frac{\partial v}{\partial t} = t q_2 \left[\gamma_{2k} \frac{dL_{2k}^c}{dt} + \sum_i \gamma_{1i} \frac{dL_{2i}}{dt} \right] + q_1 \gamma_{1k} A \frac{dL_{1k}^c}{dt} \tag{8.27}$$

where $\frac{dL_{1k}^c}{dt}$ is the expression in braces in the second term on the RHS of (8.26), that is, the tax-induced change in θ_1's screening loan size. At zero taxes, $t = 0$ equation (8.27) reduces to

$$\frac{\partial v}{\partial t} = q_1 \gamma_{1k} A \frac{dL_{1k}^c}{dt} \tag{8.28}$$

At zero taxes a marginal change in government revenue is possible while maintaining household utilities at the pre-intervention level. Hence a Pareto improvement in possible. This is because $A \neq 0$ and A is θ_1's MRS between R and L (see equation (7.5)) and would be zero if θ_{1k} were receiving a loan on his demand curve at the point of tangency to the bank's MRS at R_1.

The optimal "tax" to effect a Pareto improvement is found by solving (8.27) (noting that $\frac{dv}{dt} = 0$ is required for optimality) for t:

$$t = \frac{- q_1 \gamma_{1k} A \frac{dL_{1k}^c}{dt}}{q_2 \left[\gamma_{2k} \frac{dL_{2k}^c}{dt} + \sum_i \gamma_{2i} \frac{dL_{2i}}{dt} \right]} \tag{8.29}$$

The denominator of (8.29) times t can be interpreted as the marginal deadweight loss from the 'tax': the change in θ_2 loan sizes from their first-best levels. $q_1\gamma_{1k}A \frac{dL^c_{1k}}{dt}$ is the marginal benefit of the tax. Equality between the marginal loss and gain being required for tax optimality. t will be a subsidy, as has been conjectured, if $\frac{dL^c_{1k}}{dt}$ is positive. From equations (8.11) to (8.15) $\frac{dL^c_{1k}}{dt} > 0$ is implied by $\frac{dR_2}{dt} < 0$. One way to interpret such a subsidy in practical terms is to consider tax concessions on interest rate payments, a progressive tax rebate reduces the effective interest rate paid by those charged the highest loan rates.

8.5: Conclusion

We have shown that a Pareto improvement can be brought about by a subsidy on the highest interest rate in markets where banks are screening by loan size. The direction of this intervention (that is, a subsidy rather than a tax) is dictated by the need to increase the loan size of the signalling borrowers subject to the requirement that contracts remain incentive compatible. The use of loan size as the screening variable implies that the signalling borrowers are constrained to be off their loan demand curves. Hence MRS differ, production is not optimal and price changes have effects on signallers' profit functions which cannot be eliminated by the envelope theorem. The CE with screening can therefore be improved upon.

It should be noted that the subsidy derived here is a policy specific to the model used, particularly the definition of risk used in Chapter 7. If we had assumed that higher risk borrowers had lower expected returns, intuition and the model comparisons of Chapter 7 suggest that a tax would be the appropriate intervention. Greenwald and Stiglitz were making the general point that a planner can improve upon a competitive equilibrium whenever markets are missing and information is imperfect. The use of their methodology has permitted us to show the type of policy required in a particular credit market, as well as to gain some intuition as to underlying reasons policies can improve efficiency since the non-equality of MRS is a feature common to screening models. The feasibility of the subsidy could be questioned. However, its use requires no more information than the banks have for screening.

Chapter 9: Summary of Principal Results and Conclusions

This thesis had two points of departure. First, was the perception that a re-examination of the McKinnon and Shaw hypotheses was called for by theoretical advances in the economics of information and the strategic analysis of firms. Second, we wished to test whether informal hypotheses about bank behaviour in the Caribbean made economic sense, that is, could they be derived as noncooperative equilibria of games which were a reasonable representation of the institutional structure. The two points have proved highly complementary in that we have been able to use models based on Caribbean conditions to test not only hypotheses about those countries but some of the McKinnon and Shaw tenets. Those considered were: that liberalized banks would raise deposit rates to levels reflecting scarce capital, that liberal charter policy would result in competition-enforcing entry, and that, in the absence of interest rate controls, banks could allocate credit to its best uses without rationing. In each case we found a model whose noncooperative equilibrium failed to confirm the outcomes postulated by McKinnon and Shaw, although we introduced no assumptions which violated theirs. Indeed, our information assumptions could be considered representative of their informal descriptions of fundamentals. In two instances we were also able to show, contrary to liberalization ideas, that there existed a government regulation that could improve on the market equilibrium. However, we have simultaneously found support for the more fundamental of the McKinnon and Shaw arguments - subtleties which have hitherto largely been ignored from the policy and operational viewpoints. McKinnon and Shaw pointed out that financial intermediaries are producers and carriers of information in economies with poor and costly information flows. Each of our models showed that there are feasible arrangements by which banks can acquire information about their customers.

In Chapter 4 we showed that in the long-run equilibrium of a search market, tacit collusion can be supported as a noncooperative equilibrium since banks who are similar will set the monopsony rate. With switching costs it is in the individual interest of each bank to bid down the deposit rate: the incentive to compete is reduced by the knowledge that small rate increases will not lead to an expansion in supply. This result shows that unrestricted markets need not achieve

an efficient result, and explained observations in the Caribbean. Of more general interest was our derivation of switching cost as the interest rate premium required by long-term customers to switch banks. We showed that searching customers uncertain *ex ante* of the service offered by banks would develop switching costs once a satisfactory bank had been found, as a result of the enhanced service quality provided once they had established a reputation with the bank. The idea that switching costs are connected with uncertainty about future gain has been suggested by writers like Schmalensee [1982] and Klemperer [1987a]. However, its application did not differ substantially from transactions or search costs in that switching costs were modelled only as a cost of leaving a known supplier. We argued that switching costs place greater restrictions on mobility because they reflect not only the cost of acquiring information about a new supplier, but also foregoing the gain from remaining with the current bank for improved service.

In Chapter 5 we argued that this gain in effect enforces an implicit contract between bank and customer. The bank wants to keep the client in order to learn his needs and reliability and can do so by offering him a reward contingent on behaviour. Since this reward itself entails expected profits for the bank, the client considers the promise credible. This chapter argues that the models which derive tenure payments as rewarding effort and conveying information are usefully applied to the banking market and gives an example to illustrate this. This explanation of switching costs may help us to understand why rational individuals enter relations they know may entail later opportunistic exploitation.

Another aspect of the sluggish market share created by switching costs was examined in Chapter 6, viz. the difficulties it may create for potential market entrants. We considered the case of (mainly) multinational banks who are aware of switching costs but see expansion opportunities, which we modelled as the entry of new customers. However, the exact value of such opportunities is uncertain *ex ante*. If their value proves low the new bank's failure to attract old loyal customers means that it exits because the business it does obtain provides insufficient revenue to cover its opportunity costs. Losses incurred in this process are in effect a cost of information acquisition about new markets. Three points emerged from this model. First, it provides a new example of how switching costs may influence potential competition. Switching costs have

previously been modelled as a source of entry deterrence, of 'price wars' post-entry and of col-lusive behaviour between entrant and incumbent. Here they are a source of exit. Second, we showed that in conditions of imperfect information about market prospects, freedom from charter restrictions need not be sufficient to bring about entry which improves competition. Third, we have explained observation of entry-then-exit in the Caribbean in a manner consistent with the general observation of bank-client attachments.

In Chapter 7 and 8 we moved to consideration of credit allocation and its welfare implica-tions. This is the single area of information economics applied in the financial liberalization con-text (as in the analyses of Cho [1986], Carter [1988] and McKinnon [1988]) but these applica-tions only treated Stiglitz-Weiss credit rationing. Financial liberalization requires that banks gen-erate information enabling them to place loanable funds attracted by higher deposit rates in the most efficient uses. We show that self-selection models in the credit market confirm McKinnon and Shaw's intuition about information-generating banks. We modelled an economy in which banks are initially unable to identify risk in a particular sector, as is often assumed to be true of non-traditional industries in LDCs. Identification is accomplished by the offer of contracts such that the less risky reveal themselves by the choice of lower loan size than would be demanded at the contract rate of interest. These contracts were shown to be a sequential equilibrium if the difference in marginal rates of substitution between borrowers are sufficiently large, or if the pro-portion of low risk is sufficiently low. This equilibrium improves on the use usual in credit market screening models which usually employ reactive equilibrium, violating the Nash assump-tion, or use the Nash equilibrium which is inappropriate in a sequential game. This model demonstrated that even with bank information acquisition there can be outcomes that look like rationing. Information is only acquired because marginal rates of substitution are not equated.

The welfare issue is taken up in Chapter 8. We use a simple general equilibrium model to show that government intervention can accomplish a Pareto improvement because the government's coercive power allows it to change prices, and the price change affects welfare because marginal rates of substitution are not equated. The separation achieved by contracts is illusory in welfare terms because types are tied together through the incentive compatibility

constraints. This illustrates the Greenwald-Stiglitz [1986] point that economies with incomplete markets and imperfect information are almost never constrained Pareto efficient. It is thought that this is the first specific application of the Greenwald-Stiglitz methodology to the credit market.

We believe that there are three points of general interest for financial policy in LDCs suggested by the thesis. First, once we recognize an environment with imperfect information and strategic behaviour, predictions about empirical outcomes are not clear cut. They depend on the institutional structure (which may be endogenous in the long-run), preferences and hence cultural norms, technological conditions etc. Quite detailed knowledge of particular economies may be required for reasonable predictions. Second, this is important for policy because if we cannot describe the behaviour underlying our observations, correct policies cannot be formulated. Third, there are in general government interventions which can improve on the decentralized equilibria of economies with incomplete markets and imperfect information. Since lagging economies are, almost by definition, such economies, Pareto-improving policies will almost always exist there. The problem is to find appropriate policies which do not destroy private arrangements and incentives, taking account of the fact that no intervention may be better than unsuitable intervention. While we have found policies for two of our specifications, each economy would require specific analysis.

Bibliography

Aghion,P. and Bolton,P. [1987] "Contracts as a Barrier to Entry," *American Economic Review*, 72(3), pp.388-401

Akerlof,G. [1970] "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, v.84, pp.288-300

Allen,B. [1986] "General Equilibrium with Rational Expectations," Chapter 1 in W. Hildenbrand and A. Mas-Colell (eds), *Contributions to Mathematical Economics*, North-Holland

Arrow,K. and Hahn,F. [1971] *General Competitive Analysis*, Elsevier-North Holland

Arrow,K. [1986] "Agency and the Market," Chap.23 in K.J.Arrow and M.D.Intriligator (eds.), *Handbook of Mathematical Economics,*, v.III, Elsevier-North Holland

Barnett,C. [1982] "Commercial Banking in Belize, 1960-1980," *Social and Economic Studies*, 31(4), pp.136-169

Baron,D. and Besanko,D. [1984] "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, v.1, pp.267-302

Baum,D. [1974] *The Banks of Canada in the Commonwealth Caribbean*, Praeger

Baumol,W. [1982] "Contestable Markets: an Uprising in the Theory of Industrial Structure," *American Economic Review*, 72(1), pp.1-15

Becker,G. [1962] "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, v.70, pp.9-49

Benoit,J.-P. [1984] "Financially constrained Entry in a Game with Incomplete Information," *Rand Journal of Economics*, 15(4), pp.490-499

Bernanke,B. and Gertler,M. [1987] "Banking and Macroeconomic Equilibrium," Chapter 5 in W.Barnett and K.Singleton (eds.), *New Approaches to Monetary Theory*, Cambridge University Press

Besanko,D. [1985] "Multiperiod Contracts between Principal and Agent with Adverse Selection," *Economics Letters*, v.17, pp.33-37

Besanko,D. and Thakor,A. [1987a] "Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets," *International Economic Review*, v.28, pp.671-689

_____ [1987b] "Competitive Equilibrium in the Credit Market under Asymmetric Information," *Journal of Economic Theory*, v.42, pp.168-182

Bester,H. [1985] "Screening vs. Rationing in Credit Markets with Imperfect Information," *American Economic Review*, v.75, pp.850-855

_____ [1987] "The Role of Collateral in Credit Markets with Imperfect Information," *European Economic Review*, v.31, pp.887-899

Bhattacharya,S. and Gale,D. [1987] "Preference shocks, liquidity and central bank policy," Chapter 4 in W.Barnett and K.Singleton (eds.), *New Approaches to Monetary Theory*, Cambridge University Press

Blackman,C. [1985] "How a Small Country Banker views the International Financial Outlook," Address to the Tenth World Banking Conference, 1984, reprinted in Central Bank of Barbados, *Economic Review*, 12(1), pp.34-38

Blackwell,N. and Santomero,A. [1982] "Bank Credit Rationing and the Customer Relation," *Journal of Monetary Economics*, 9(1), pp.121-129

Blinder,A. [1987] "Credit Rationing and Effective Supply Failure," *Economic Journal*, v.97, pp.327-352

Bourne,C. [1982] "Notes on Financial Changes in Trinidad and Tobago,1966-78", *Social and Economic Studies*, 31(4), pp.171-191

_____ [1984] "Structure and Performance of Commercial Banking in Trinidad and Tobago," Paper presented at Sixteenth Annual Conference, Regional Programme of Monetary Studies

_____ [1985] "Banking in Boom and Bust Economies: Lessons from Trinidad and Tobago and Jamaica," *Social and Economic Studies*, 34(4), pp.139-162

Boyd,J. and Prescott,E. [1986] "Financial Intermediary-Coalitions," *Journal of Economic Theory*, v.38, pp.211-232

Bull,C. [1987] "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, v.102, pp.147-159

Bulow,J., Geanakoplos,J. and Klemperer,P. [1985] "Multimarket Oligopoly: Strategic Substitutes and Complements," *Journal of Political Economy*, 93(3), pp.488-511

Carter,M. [1988] "Equilibrium Credit Rationing of Small Farm Agriculture," *Journal of Development Economics*, v.28, pp.83-103

Central Bank of Barbados [1986] *A Regulatory Framework for Financial Institutions in Barbados*, Draft Discussion (Green) Paper

_____ [1988] *Annual Statistical Digest*

Cho,I-K. and Kreps,D. [1987] "Signalling Games and Stable Equilbria," *Quarterly Journal of Economics*, 102(2), pp.179-222

Cho,Y. [1986] "Inefficiencies from Financial Liberalization in the Absence of Well-Functioning Equity Markets," *Journal of Money, Credit and Banking*, 18(2), pp.191--199

Clemenz,G. [1986] *Credit Markets with Asymmetric Information*, Springer-Verlag Lecture Notes in Economics and Mathematical Systems, No.272

Coase,R. [1960] "The Problem of Social Cost," *Journal of Law and Economics*, v.3, pp.1-44

Conrad,C. [1983] "The Advantage of Being First and Competition Between Firms," *International Journal of Industrial Organization," v.1, pp.353-364*

Cooter,R. [1987] "The Coase Theorem" in *The New Palgrave*

Debreu,G. [1959] *The Theory of Value: an Axiomatic Analysis of Economic Equilibrium*, Yale University Press

de Meza,D. and Webb.D. [1987] "Too Much Investment: A Problem of Asymmetric Information," *Quarterly Journal of Economics*, v.102, pp.281-292

_____ [1988] "Credit Market Efficiency and Tax Policy in the Presence of Screening Costs," *Journal of Public Economics*, v.36, pp.1-22

Devinney,T. [1986] *Rationing in a Theory of the Banking Firm*, Berlin: Springer-Verlag Studies in Contemporary Economics

Diamond,D. and Dybvig,P. [1983] "Bank Runs, Deposit Insurance and Liquidity," *Journal of Political Economy*, 91(3), pp.401-419

Diamond,D. [1984] "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, v.51, pp.393-414

Eatwell,J. Milgate,M. and Newman,P. (eds.) [1987] *The New Palgrave: A New Dictionary of Economics*, Macmillan Press Limited

Engers,M. [1987] "Signalling with many Signals," *Econometrica*, 55(3), pp.663-674

Farrell,J. [1986] "A Note on Inertia in Market Share," *Economics Letters*, v.21, pp. 73-75

Farrell,J. and Shapiro,C. [1988] "Dynamic Competition with Switching Costs," *Rand Journal of Economics, 19(1), pp.123-137*

_____ [1989] "Optimal Contracts with Lock-In," *American Economic Review*, 79(1), pp.51-68

Flannery,M. [1982] "Retail bank deposits as Quasi-fixed factors of production," *American Economic Review* , 72(3), pp.527-536

Fried,J. and Howitt,P. [1980] "Credit Rationing and Implicit Contract Theory," *Journal of Money, Credit and Banking*, 12(3), pp.471-487

Friedman,J. [1986] *Game Theory with Applications to Economics*, Oxford University Press

Fudenberg,D. and Tirole,J. [1983] "Learning-by-doing and market performance," *Bell Journal of Economics*, v.14, pp.522-531

_____ [1984] "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look," *American Economic Review, Papers and Proceedings,* 74(2), pp.361-366

_____ [1986] "A Theory of Exit in Duopoly," *Econometrica,* 54(4), pp.943-960

Galbis,V. [1979] "Structuralism and Financial Liberalization," *Finance and Development* ,13(2), pp.33-37

Gale,D. [1982] *Money: in equilibrium,* Nisbet/Cambridge

_____ [1987a] "Price Setting and Competition in a simple duopoly model," mimeo

_____ [1987b] "A Walrasian Theory of Markets with Adverse Selection, with an application to credit markets," mimeo

_____ and M.Hellwig [1985] "Incentive Compatible Debt Contracts: the One-Period Problem," *Review of Economic Studies,* v.52, pp.647-663

Geanakoplos,J. and Polemarchakis,H. [1986] "Existence, regularity and constrained suboptimality of competitive portfolio allocations when the asset market is incomplete," pp.65-95 in W.Heller, R.Ross and D.Starrett (eds), *Uncertainty, Information and Communication: Essays in Honour of Kenneth J. Arrow, Vol.III*

Geroski,P., Phlips,L. and Ulph,A. (eds.) [1985] *Oligopoly, Competition and Welfare,* Basil Blackwell in cooperation with the Journal of Industrial Economics

Ghatak,S. [1981] *Monetary Economics in Developing Countries,* Macmillan

Ghemawat,P. and Nalebuff,B. [1985] "Exit," *Rand Journal of Economics,* 16(2), pp.184-194

Gilbert,R. [1988] "Mobility Barriers and the Value of Incumbency," in R.Schmalensee and R.Willig (eds.) *Handbook of Industrial Organization,* Elsevier-North Holland

_____ and Vives,X. [1986] "Entry Deterrence and the Free Rider Problem," *Review of Economic Studies,* v.53, pp.71-83

Goodhart,C. [1984] *Monetary Theory and Practice: the U.K. Experience,* Macmillan Press

Greenwald,B., Stiglitz,J. and Weiss,A. [1984] "Informational Imperfections in the Capital Markets and Macroeconomic Fluctuations," *American Economic Review, Papers and Proceedings,* v.74, pp.194-200

Greenwald,B. and Stiglitz,J. [1986] "Externalities in Economies with Imperfect Information and Incomplete Markets," *Quarterly Journal of Economics,* 101(2), pp.229-264

Grossman,S. [1976] "On the Efficiency of Competitive Stock Markets where Traders have Diverse Information," *Journal of Finance,* 31(1), pp.573-585

_____ [1977] "The Existence of Futures Markets, Noisy Rational Expectations and Informational Externalities," *Review of Economic Studies*, v.44, pp.431-449

_____ [1981] "An Introduction to the Theory of Rational Expectations under Asymmetric Information," *Review of Economic Studies*, v.48, pp.541-559

_____ and J. Stiglitz [1980] "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70(3), pp.393-408

Grout,P. [1984] "Investment and Wages in the absence of Binding Contracts: A Nash Bargaining Approach," *Econometrica*, v.52, pp.449-460

Guasch,J. and Weiss,A. [1980] "Wages as Sorting Mechanisms in Competitive Markets with Asymmetric Information: A Theory of Testing," *Review of Economic Studies*, v.47, pp.653-664

Hadley,G. [1964] *Nonlinear and Dynamic Programming*, Addison-Wesley Publishing Co. Inc

Hahn,F. [1988] "On Monetary Theory," *Economic Journal*, v.98, pp.957-973

Harris,L. [1985] "Financial Reform and Economic Growth: A New Interpretation of South Korea's Experience," mimeo, The Open University

Hart,O. [1975] "On the Optimality of Equilibrium when the Market Structure is Incomplete," *Journal of Economic Theory*, v.11, pp.418-443

_____ [1986] Comment on Stiglitz and Weiss [1986], pp.136-143

_____ and Holmstrom,B. [1987] "The Theory of Contracts," Chapter 3 in T.Bewley (ed.), *Advances in Economic Theory: Fifth World Congress*, Cambridge University Press

Hayek,F. [1945] "The Use of Knowledge in Society," *American Economic Review*, v.35, pp.519-530

Hellwig,M. [1987] "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection," *European Economic Review*, v.31, pp.319-325

Holmstrom,B. [1983] "Equilibrium Long-Term Contracts," *Quarterly Journal of Economics, Supplement*, v.98, pp.23-54

Holmstrom,B. and Myerson,R. [1983] "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, 51(6), pp. 1799-1819

_____ and Weiss,L. [1985] "Managerial Incentives, Investment and Aggregate Implications: Scale Effects," *Review of Economic Studies*, v.52, pp.403-425

Howard,M. [1976] "Interest Rate Behaviour in an Open Developing Economy: the Barbadian Experience, 1965-76," Central Bank of Barbados *Quarterly Report*, 3(4), pp.24-50

Ireland,N. [1987] *Product Differentiation and Non-price Competition*, Basil Blackwell

Jacklin,C. and Bhattacharya,S. [1988] "Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications," *Journal of Political Economy*, 96(3), pp.568-592

Jaffee,D. and Russell,T. [1976] "Imperfect Information, Uncertainty and Credit Rationing," *Quarterly Journal of Economics*, v.90, pp.651-660

Klemperer,P. [1986a] "The Welfare Effects of New Entry," Chapter 5 in Markets with Switching costs, Ph.D. dissertation, Stanford University Graduate School of Business

_____ [1986b] "Price Wars caused by New Entry," Chapter 6 in Markets with Switching Costs, Ph.D. dissertation, Stanford University Graduate School of Business

_____ [1987a] "Markets with Consumer Switching Costs," *Quarterly Journal of Economics*, v.102, pp.375-394

_____ [1987b] "Entry Deterrence in Markets with Consumer Switching Costs," *Economic Journal Conference Papers*, v.97, pp.99-117

_____ [1987c] "The Competitiveness of Markets with Consumer Switching Costs," *Rand Journal of Economics*, v.18, pp.138-150

Kohn,M. and Shavell,S. [1974] "The Theory of Search," *Journal of Economic Theory*, v.9, pp.93-123

Kreps,D. and Wilson,R. [1982] "Reputation and Imperfect Information," *Journal of Economic Theory*, v.27, pp.253-279

Kreps,D. and Wilson,R. [1982] "Sequential Equilibria," *Econometrica*, 50(4), pp.863-894

Laffont,J-J. and Maskin,E. [1982] "The Theory of Incentives: an overview," in W. Hildenbrand (ed.), *Advances in Economic Theory*, Cambridge University Press

Laffont,J-J. [1985] "On the Welfare Analysis of Rational Expectations Equilibria with Asymmetric Information," *Econometrica*, 53(1), pp.1-29

_____ and Tirole,J. [1988] "The Dynamics of Incentive Contracts," *Econometrica*, 56(5), pp.1153-1175

Lambert,R. [1983] "Long-term Contracts and Moral Hazard," *Bell Journal of Economics*, 14(2), pp.441-452

Larson,R. and Casti,J. [1978] *Principles of Dynamic Programming: Part I*, Marcel Dekker Inc.

Lazear,E. [1981] "Agency, Earnings Profiles, Productivity and Hours Restrictions," *American Economic Review*, v.71, pp.606-620

Leland,H. and Pyle,D. [1977] "Informational Asymmetries, Financial Structure and Financial Intermediation," *Journal of Finance*, 82(2), pp.371-388

Levhari,D. and Srinivasan,T. [1969] "Optimal Savings under Uncertainty," *Review of Economic Studies*, V.36, pp.153-163

Lewis,W.A. [1949] "The Economics of Loyalty," Chapter III in *Overhead Costs*, London: Unwin University Books

Lippman,S. and McCall,J. [1976] "The Economics of Job Search: A Survey," *Economic Inquiry*, v.14, Part I, pp.155-189

Luben,M. [1983] "The Financial System in Belize," mimeo, paper presented at the Fifteenth Annual Regional Programme of Monetary Studies Conference

Lucas,R. [1972] "Expectations and the Neutrality of Money," *Journal of Economic Theory*, v.4, pp.103-124

MacLeod,B. and Malcolmson,J. [1989] "Implicit Contracts, Incentive Compatility and Involuntary Unemployment," *Econometrica*, v.57, pp.312-322

Mankiw,N. [1986] "The Allocation of Credit and Financial Collapse," *Quarterly Journal of Economics*, v.101, pp.455-470

McClean,W. [1975] *Money and Banking in the East Caribbean Currency Area*, Mona, Jamaica: Institute of Social and Economic Research, University of the West Indies

McKinnon,R. [1973] *Money and Capital in Economic Development*, Washington,D.C.:The Brookings Institution

_____ [1988] "Financial Liberalization in Retrospect: Interest Rate Policies in LDCs," Chapter 13 in G. Ranis and T. Schultz (eds.), *The State of Development Economics*, Basil Blackwell

Merton,R. [1969] "Lifetime Portfolio Selection under Uncertainty: the continuous case," *Review of Economics and Statistics* , V.51, pp.247-257

Milde,H. and Riley,J. [1988] "Signalling in Credit Markets," *Quarterly Journal of Economics*, 103(1), pp.101-129

Milgrom,P. and Roberts,J. [1982] "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, v.27, pp.280-312

Mirrlees,J. [1974] "Optimal Accumulation under Uncertainty: the case of stationary returns to investment", in J. Dreze (ed.) *Allocation under Uncertainty: Equilibrium and Optimality*, London: Macmillan, pp.36-50

Mortensen,D. [1986] "Closed Form Equilibrium Price distribution functions," mimeo

Myint,H. [1971] "Dualism and the Internal Integration of the Underdeveloped Economies," Chapter 14 in *Economic Theory and the Underdeveloped Countries*, Oxford University Press

Ordover,J. and Weiss,A. [1981] "Information and the Law: Evaluating Legal Restrictions on Competitive Contracts," *American Economic Review, Papers and Proceedings,* v.81, pp.399-404

Ostroy,J. [1989] "Money and general equilibrium theory," in *The New Palgrave*

Parris,C. [1984] "Political Economy of Financial Collapse: the Saga of International Trust Ltd., Trinidad and Tobago, 1982/84," Paper presented at the Sixteenth Annual Conference of the Regional Programme of Monetary Studies

Park,Y. [1988] "Korea," Chap. 15 in R.Dornbusch and F.Helmers (eds.), *The Open Economy: Tools for Policymakers in Developing Countries,* Oxford University Press

Phelps,E. [1962] "The accumulation of risky capital: a sequential utility analysis," *Econometrica,* 30(4), pp.729-743

Postlewaite,A. and Vives,X. [1987] "Bank Runs as an Equilibrium Phenomenon," *Journal of Political Economy,* 95(3), pp.485-491

Quinzii,M. [1988] Lecture 4: Production and the Behaviour of the firm, pp.32-39 in Lecture Notes in Incomplete Markets, D.P. A-192, Rheinische Friedrich-Wilhelms, University of Bonn

Radner,R. [1972] "Existence of equilibrium of plans, prices and price expectations in a sequence of markets," *Econometrica,* v.40, pp.289-304

_____ [1981] "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship," *Econometrica,* v.49, pp.1127-1148

_____ [1982a] "Equilibrium under Uncertainty," Chapter 20 in K.Arrow and M.Intriligator (eds.) *Handbook of Mathematical Economics,* Vol. II, Elsevier-North Holland

_____ [1982b] "The role of private information in markets and other organizations," Chapter 3 in W. Hildenbrand (ed.) *Advances in Economic Theory,* Cambridge University Press

Radner,R. and Stiglitz,J. [1984] "A Nonconcavity in the Value of Information," Chapter 3 in M. Boyer and R. Kihlstrom (eds.) *Bayesian Models in Economic Theory,* Elsevier Science Publishers B.V.

Ramkissoon,R. [1981] "Commercial Bank Portfolio Behaviour in Trinidad and Tobago," mimeo, Paper presented at Regional Monetary Studies Conference

Ramkrishnan,R. and Thakor,A. [1984] "Informational Reliability and a Theory of Financial Intermediation," *Review of Economic Studies,* v.51, pp.415-432

Ramsaran,R. [1984] *The Monetary and Financial System of the Bahamas, Growth, Structure and Operation,* Institute of Social and Economic Research, University of the West Indies, Jamaica

Riley,J. [1979] "Informational Equilibrium," *Econometrica*, 47(2), pp.331-359

_____ [1985] "Competition with Hidden Knowledge," *Journal of Political Economy*, 93(5), pp.958-976

_____ [1987] "Credit Rationing: A Further Remark," *American Economic Review*, 77(1), pp.224-227

Roberts,J. [1986] "A Signalling Model of Predatory Pricing," pp.75-93 in D.Morris, P.Sinclair, M.Slater and J.Vickers (eds.) *Strategic Behaviour and Industrial Competition*, Oxford: Clarendon Press

_____ [1987] "Battles for Market Share: incomplete information, aggressive strategic pricing, and competitive dynamics," Chapter 4 in T.Bewley (ed.) *Advances in Economic Theory - Fifth World Congress*, Cambridge University Press

Roberts,K. [1982] "Long-Term Contracts," mimeo, University of Wartwick

Rosenthal,R. [1982] "A Dynamic Model of Duopoly with Customer Loyalties," *Journal of Economic Theory," v.27, pp.69-76*

Rothschild,M. and Stiglitz,J. [1970] "Increasing Risk:I. A Definition," *Journal of Economic Theory*, v.2, pp.225-243

_____ [1976] "Equilibrium in Competitive Insurance Markets: and Essay in the Economics of Imperfect Information," *Quarterly Journal of Economics*, v.90, pp.629-649

Salop,J. and Salop,S. [1976] "Self-Selection in the Labour Market," *Quarterly Journal of Economics,FR v.90, pp.619-627*

Salop,S. [1977] Comments on J.G. Greenhut, M.L. Greenhut and W.H. Kelly, "A Spatial-Theoretical Perspective for Bank Merger Regulations," in Federal Reserve Bank of Chicago, *Proceedings of a Conference on Bank Structure and Performance*, pp.255-259

_____ and Stiglitz,J. [1977] "Bargains and Ripoffs: A Model of monopolistically competitive price dispersion," *Review of Economic Studies*, v.47, pp.493-510

_____ [1982] "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents," *American Economic Review*, 72(5), pp.1121-1130

Samuelson,P. [1969] "Lifetime portfolio selection by dynamic stochastic programming," *Review of Economics and Statistics*, V.51, pp.239-246

Santomero,A. [1984] "Modelling the Banking Firm," *Journal of Money, Credit and Banking* 16(4)

Scherer,F. [1980] *Industrial Market Structure and Economic Performance*, Chicago: Rand McNally College Publishing Company

Schmalensee,R. [1982] "Product Differentiation Advantages of Pioneering Brands," *American Economic Review*, 72(3), pp.349-365

Selten,R. [1975] "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, v.4, pp.25-55

Shapiro,C. [1982] "Consumer Information, Product Quality and Seller Reputation," *Bell Journal of Economics*, 13(1), pp.20-35

Shaw,E. [1973] *Financial Deepening in Economic Development*, Oxford University Press

Shubik,M. [1987] "A Game Theoretic Approach to the Theory of Money and and Financial Institutions," European University Institute, Department of Economics, W.P. No. 87/280

Stiglitz,J. [1984] "Information and Economic Analysis: a perspective," *Economic Journal*, Conference Papers, pp.21-41

_____ [1987] "Competitivity and the number of firms in market: Are Duopolies more competitive than Atomistic Markets," *Journal of Political Economy*, 95(5), pp.1041-1061

Stiglitz,J. and Weiss,A. [1981] "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71(3), pp.393-410

_____ [1983a] "Incentive Effects of Terminations: Applications to the Credit and Labour Markets," *American Economic Review*, v.73, pp.912-927

_____ [1983b] "Sorting out the Differences between Screening and Signalling Models," Columbia University, Department of Economics Discussion Paper No. 224

_____ [1986] "Credit Rationing and Collateral," in J.Edwards, J.Franks, C.Mayer and S.Schaefer (eds.), *Recent Developments in Corporate Finance*, Cambridge University Press

_____ [1987] "Credit Rationing: Reply," *American Economic Review*, 77(1), pp.228-231

Thomas,J. and Worrall,T. [1988] "Self-Enforcing Wage Contracts," *Review of Economic Studies*, v.55, pp.541-554

Van Wijnbergen,S. [1983] "Interest Rate Management in LDCs," *Journal of Monetary Economics*, 12(3), pp.433-452

von Weizsacker,C. [1980] "A Welfare Analysis of Barriers to Entry," *Bell Journal of Economics*, 11(2), pp.399-420

_____ [1984] "The Costs of Substitution," *Econometrica*, 52(5), pp.1085-1116

Wette,H. [1983] "Collateral in Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 73(3), pp.442-445

Williamson,S. [1986] "Costly Monitoring, Financial Intermediation and Equilibrium Credit Rationing," *Journal of Monetary Economics*, v.18, pp.159-179

_____ [1987] "Costly Monitoring, Optimal Contracts and Equilibrium Credit Rationing," *Quarterly Journal of Economics*, v.102, pp.135-145

_____ [1988] "Liquidity, banking and bank failure," *International Economic Review*, 29(1), pp.25-44

Wilson,C. [1977] "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory*, v.16, pp.167-207

Wilson,R. [1989] "Entry and Exit," Chapter 8 in G.Feiwel (ed.) *The Economics of Imperfect Competition and Employment: Joan Robinson and Beyond*, Macmillan

Wolinsky,A. [1983] "Prices as Signals of Product Quality," *Review of Economic Studies*, V.50, pp647-658

Wood,J. [1974] "A Model of Commercial Bank Loan and Investment Behaviour," in H.G.Johnson and A.R.Nobay (eds), *Issues in Monetary Economics*, Oxford University Press

World Bank,The [1989] *World Development Report 1989*, Oxford University Press

Worrell,D. and Prescod,R. [1983] "The Development of the Financial Sector in Barbados, 1946-1982," Central Bank of Barbados *Economic Review*, 10(2), pp.9-26

Worrell,D. [1985] "Monetary Mechanisms in Open Economies: A Model for the Caribbean," Chapter 6 in M.Connolly and J.McDermott (eds.) *The Economics of the Caribbean Basin*, N.Y.: Praeger

_____ [1987] *Small Island Economies*, N.Y.: Praeger

Yanelle,M. [1986] Bertrand Competition among intermediaries, D.P. A-31, Rheinische Friedrich-Wilhelms, University of Bonn