

CONTRACTUAL CONSTRAINTS ON FIRM PERFORMANCE IN DEVELOPING COUNTRIES¹

Dilip Mookherjee²

July 1999

1 Introduction

In this paper I will be concerned with distinctive characteristics of firms and industry structures in developing countries, stemming from asymmetric information and related contracting impediments. These constraints are also frequently referred to as ‘institutional’ or as embodying ‘transaction costs’. The role of such institutions has been discussed extensively in traditional literature on the subject, consisting mainly of descriptive and historical accounts. However, as T.W. Schultz noted three decades ago: “in thinking about institutions, the analytical cupboard is bare...” (Schultz(1968)). In contrast the recent literature is comprised of a profusion of analytical ideas, though empirical tests of some of these have begun to emerge very recently. My intention will be to focus selectively on a few strands of this literature that I am familiar with and find exciting. In particular I will try to concentrate on themes which have been close to my own recent research, and admit a coherent analytical framework.

The term ‘contracting impediments’ refers to a variety of market imperfections that firms in developing countries are especially prone to, that are sought to be overcome by

¹Prepared for the Session on Industrial Organization and Development at the Latin American meetings of the Econometric Society, August 1999. I am grateful to Abhijit Banerjee, Pranab Bardhan, Hadi Esfahani, Karla Hoff, Rachel Kranton and Debraj Ray for many conversations and useful suggestions.

²Department of Economics, Boston University; dilipm@bu.edu

various non-market institutions. These include distinct contracting arrangements and organizational forms, formal institutions of the state, informal institutions such as family and community networks, and related social norms. I will avoid focusing on the role of the state, despite its historical role in shaping industry performance in developing countries. Indeed, a very large literature has traditionally been devoted to this topic both from a positive and normative standpoint: infrastructure provision, financial intermediation, trade and industrial policies, and the performance of public sector firms. I wish to focus instead on organizational innovations and informal institutions that have spontaneously evolved in the absence of formal state support, partly because they have not received adequate emphasis in traditional literature. Moreover, these are likely to achieve greater significance in the post-1990s environment of substantive deregulation and ‘liberalization’ of free market forces.

I will argue that the diverse problems of information and trust encountered in running a business in different markets, industries and countries can be understood as stemming from a common problem of moral hazard, based on unobservability of certain dimensions of ‘effort’, that is compounded by limitations on contracting structures imposed by financial constraints and/or weaknesses in contract enforcement mechanisms. Moral hazard refers to scope for *ex post* opportunistic behavior. It represents problems of lack of trust and reliability, that are typically overcome via explicit contracting or formal governance structures in more developed economies. Developing societies are more prone to these problems either because of greater limitations on the design of contracts owing to financial constraints (that have greater bite owing to more significant capital market imperfections), or because of weaker contract enforcement institutions (resulting from weaknesses in monitoring and information systems, and weaker legal institutions). Owing to either of these constraints, it is more difficult to discipline dysfunctional behavior via appropriately structured punishments that are administered by formal institutions. In particular, it will be argued that the appropriate model of the moral hazard problem is not the classical one associated with the work of Mirrlees (1975), Holmstrom (1979), or Grossman and Hart (1983), but one subject to additional financial, contracting or enforcement constraints. Consequently these problems of trust play a larger role in developing countries.

Three alternative resolutions of the compounded moral hazard problem owing to lack of effective ‘sticks’ are described in this paper. The first involves the design of contracts which use the carrot rather than stick to reward abstinence from dysfunctional behavior, e.g., the payment of ‘efficiency wages’ to workers, or ‘informational rents’ to clients. In this instance relevant performance characteristics are contractible, and it is possible to enforce contracts that are conditioned on measured performance. The only constraints are financial constraints that limit the liability of the agent.

The second way of mitigating the moral hazard problem in contexts where performance is noncontractible, or contracts with complex contingencies cannot be written or enforced, is to design forms of joint ownership that mitigate incentives for *ex post* opportunism.

The third approach, employed whenever neither of the above two approaches work, relies on reputational concerns, i.e., the threat of withdrawing future business dealings.

Illustrations of all three mechanisms will be provided in the paper, along with empirical studies that document their significance.

These contracting impediments have a number of interesting implications. They suggest the role of a variety of structural determinants of industry performance in developing countries that tend to be rarely emphasized in the mainstream industrial organization literature, such as:

- wealth and income inequality
- financial intermediation institutions
- community networks (social capital)
- labor skill availability (human capital)
- infrastructure provision, especially their quality and reliability
- relationship of agroprocessing industry with agricultural suppliers, and the nature of agrarian structure

They provide support for the notion that ‘human development’ may be desirable not only for its own sake, but can also be instrumental in fostering industrial progress. The preconditions for successful industrialization may well include agricultural prosperity; absence of a ‘feudal’ or ‘semi-feudal’ agrarian structure, extreme economic inequality or ethnic fragmentation; and the provision of reliable infrastructure services. They help explain why, for instance:

- ‘doing business’ is more difficult in developing countries, owing to greater vulnerability to informational and contracting constraints (a theme which goes back to Akerlof (1970)), thus providing a microfoundation for why capital may not flow from developed countries to developing countries to the extent predicted by a neoclassical growth model, a question raised by endogenous growth theorists (e.g., Lucas (1988))
- industrialization patterns within the developing world tend to be uneven, restricted to specific countries, locations or ethnic communities that possess favorable institutional attributes
- redistributive government policies such as land reforms and provision of universal education may have benign effects on industrial performance, and may thus have played an important role in the ‘East Asian miracle’
- successful firms tend either to be large family-controlled conglomerates, or belong to close-knit networks or industrial clusters frequently restricted to specific ethnic communities
- occasionally some ‘unusual’ organizational forms achieve striking success, such as Indian sugar cooperatives or Chinese township and village enterprises, and why these may be difficult to replicate elsewhere
- despite substantial deregulation, conglomerate business groups continue to play an important role

On the policy dimension, specific lessons are slower to emerge. Some broad lessons of a negative nature are however evident from these models. For one, informational and

contracting constraints create a setting where market activity is characterized by pervasive externalities, so there is no presumption that the ‘invisible hand’ works in a way suggested by the First Welfare Theorem. However, the argument for purposive state intervention is not evident either, since governments are equally bound by these informational and enforcement constraints, besides possible vulnerability to various forms of misgovernance (see Kahn and Mookherjee (1995) for further discussion). There is no overarching general argument for or against state intervention: in certain contexts they may play a constructive role, while in others they make matters worse. The ‘quality’ of governance in a given society is an important institutional attribute, that deserves serious study in its own right. Other policy lessons of a more specific nature include:

- the ambiguous effects of increased competition on prices and firm performance,
- the difficulties of breaking into new export markets despite the existence of underlying comparative advantage, and
- the importance of financial market institutions that facilitate intermediation and expanded firm access to external capital, as well as the provision of reliable infrastructure

The paper is organized as follows. Section 2 illustrates the contractual ‘informational rent’ resolution of the moral hazard problem, in the context of credit market imperfections, such as credit rationing and large divergences between the cost of external and internal capital. Here we describe a one-sided moral hazard model where the limits to the financial liability of agents represent the principal contracting constraint. The model explains the role of (liquid) wealth effects, and of property right allocations. An extension of the model to a dynamic setting illustrates the role of monopolistic credit market structure in precipitating low-productivity-investment traps. Empirical evidence concerning liquidity and wealth effects on contractual structure and productivity is also described.

Section 3 describes the second mode of resolution of the moral hazard problem, via appropriate design of ownership rights. Here we consider settings of two-sided moral hazard, which can motivate joint ownership structures, where multiple input suppliers jointly

own the firm. Applications to the nature of foreign direct investment in China, and the performance of Indian sugar cooperatives, are described.

Section 4 turns to reputation as a device for reining in opportunistic behavior. It helps explain why the provision of high quality tends to be more problematic in developing countries, and why they confer a special role to community contractual networks as an alternative to vertically integrated enterprises. The performance of these distinct organizational forms in the context of the Tirupur garment export industry in Southern India are subsequently discussed. Another illustration is the greater prevalence of dual labor markets in developing countries, which are argued to provide possible explanations for differences in labor productivity between developed and developing countries unaccounted by differences in capital or technology. A recent study documenting the importance of reputational factors in the Indian software export industry is also described.

Finally, Section 5 concludes with a summary of the principal implications, and possible directions for future research.

2 Credit Market Imperfections: One-Sided Moral Hazard with Financial Constraints

Most development economists are familiar with a large literature on credit rationing in the context of agrarian institutions (see for example Bardhan (1989)). Similarly the role of land property rights has been stressed in discussions of agricultural performance. The analytic issues are not fundamentally distinct in the industrial context, especially when one considers the performance of small firms. Capital requirements tend to be larger in industry, owing to the greater scope for economies of scale, so the impact of limited credit access plays a potentially larger role compared with agriculture. Imperfections in formal credit markets tend to be more pervasive in developing countries, owing to a poorly developed intermediary sector, weaknesses in information and enforcement of formal credit contracts, and thinness of bond and equity markets. These imperfections raise the cost of ‘external’ capital in ways

that differ with specific borrower characteristics, such as existing wealth and extent to which these are collateralizable.

In turn this has a number of key implications for explaining entry into the industrial sector and subsequent investment levels: (i) the existence of large ‘internal’ investible surpluses that obviate the need for external capital — surpluses that frequently originate in agricultural prosperity or other forms of ‘primitive accumulation’ *a la* Marx; (ii) the importance of affiliation in informal social networks — based on family, caste, clan, tribe or race — that fill the void created by formal credit institutions, owing to their superior position with regard to information and enforcement; and (iii) a positive correlation between internal capital and access to external capital — the phenomenon that the wealthy have better access to external credit. In particular, this implies that wealth inequality matters: the poor cannot get going because they have neither internal nor access to external capital. In addition, the effect of these capital market imperfections on other factor markets will also become obvious in due course; many of these would disappear substantially in a world of Walrasian credit markets.

The most well-known model of credit rationing is that of Stiglitz and Weiss (1981). This theory has been criticized by many authors for failing to be robust in a number of empirically relevant dimensions. These difficulties mainly stem from the particular definition of credit rationing employed by Stiglitz and Weiss, in particular on its emphasis of unequal treatment of observationally identical borrowers (where some get credit and others not at all). This phenomenon arises from adverse selection, where increased interest rates adversely affect the applicant pool, thereby limiting the willingness of banks to raise interest rates at times of credit scarcity. As a number of authors have pointed out, the differential treatment of identical borrowers could disappear if:

- (a) project sizes were divisible, for then limited funds could be spread among all applicants uniformly, resulting in smaller project sizes per applicant;
- (b) borrowers could post collateral, which would help banks screen different classes of borrowers, obviating the sorting role of the interest rate (Bester (1985));

- (c) average returns and variance of returns were positively correlated, in which case higher interest rates need not lower the rate of return earned by banks; and
- (d) there were multiple observationally distinguishable groups, in which case at most one such group would be rationed (Riley (1987))

In addition, the theory does not have any implications for the nature of wealth effects, an issue concerning which empirical evidence is available, and which has a number of wide-ranging implications, as elaborated further below.

An alternative definition of credit rationing would be based on the property that borrowers would like to borrow more at the current interest rate they are charged, but are prevented from doing so. In other words, borrowers are not operating on their Walrasian demand curve for loans, and face quantity constraints. Credit limits need not be absolute: a borrower's request for a larger loan size may be entertained, but only if it is accompanied by a higher interest rate, or a bigger collateral. In other words, it can be identified with the phenomenon of *nonlinear interest rates*, as depicted in Figure 1. Such nonlinear interest rates may also be accompanied (for given collateral) by some absolute limit on loan size.

Such forms of nonlinear interest rates can be the implication of use of loan size (besides collateral) as a screening device by banks under situations of adverse selection. For instance, borrowers with larger default risk will be willing to accept a higher interest rate in exchange for a bigger loan; conversely safe borrowers may have to settle for a smaller loan than they would have wished in order to signal their private information regarding their credit-worthiness. Similarly, it can be a method to limit the phenomenon of moral hazard, where larger loans increase 'debt overhang', thus encouraging borrowers to exert insufficient effort to prevent default.

We illustrate such a moral hazard model below. In contrast to standard models of moral hazard (e.g., Mirrlees (1975), Holmstrom (1979) or Grossman and Hart (1983)), its main distinguishing characteristic is the existence of limits on the liability of the agent. These constrain financial transfers between borrower and lender in the event that projects financed

Figure 1: NONLINEAR INTEREST RATES

are unsuccessful, leading to defaults which are costly to the lender. Borrowers face limits on credit access on grounds that in the absence of such limits, they would have insufficient *ex ante* incentive to prevent default, either by working hard on the financed projects or by selecting safe projects. *Ex post* the default is not intentional, but the outcome of lack of financial resources of the borrower owing to the failure of the project. Collateral posted by the borrower serves to relax credit limits because it enhances the borrower's incentive to avoid default, besides limiting the risk the lender is exposed to in the event of default. Models of this kind have been studied by Foster and Wan (1984), Shetty (1988), Dutta, Ray and Sengupta (1989), Newman (1995), Bowles and Gintis (1994, 1995), Hoff (1994), Hoff and Lyon (1994), Holmstrom and Tirole (1994), Legros and Newman (1996), Mookherjee (1997a, 1997b), Aghion and Bolton (1997), and Piketty (1997).

A related model concerns intentional *ex post* default, where the borrower has the resources to repay the loan, but chooses not to. Conversely, default can be the result of inability of lenders to enforce repayment of their loans, rather than limits on *ex post* liability of the borrower. The main discipline device here is the prospect of losing access to credit in the future from the borrower, i.e., out of the borrower's concern for his reputation. Such

models have been studied by Eaton and Gersovitz (1981), Coate and Ravallion (1993), Hoff and Stiglitz (1998), Ghosh and Ray (1996, 1999), among others, and are closely related to the reputational models to be described in Sections 5 and 6 below. The reputation model shares many features in common with the limited liability model. Larger loan sizes enhance ‘debt overhang’, thereby encouraging the borrower to ‘take the money and run’. Hence larger loans must be accompanied by higher interest rates and/or higher collateral; in many cases there may be an absolute limit to loan size. Moreover, in both theories external finance involves agency problems, and is thus more expensive than internal finance, i.e., liquid funds owned by the borrower — the key capital market imperfection. Owing to these similarities, we choose to illustrate only the limited liability model in this Section.

2.1 One-Sided Moral Hazard with Limited Liability Constraints

2.1.1 Theory

Suppose E is an entrepreneur who needs to borrow from lender L to finance input costs l on a given project. E has access to assets (or internal financial reserves) worth w . If these assets are liquid, there is a borrowing need if w is smaller than l . If they are not liquid, these assets can be pledged as collateral to the lender: indeed, we shall follow this interpretation. Then the credit needed from L is l at the commencement of the project. If the project is funded and operated, it realizes an uncertain return after some time lag. If the returns exceed $l(1+r)$, where r is the interest rate on the loan, then the loan is repaid at the stated terms. Otherwise E defaults on the loan, and his liability is limited to the sum of the project returns and loss of the pledged collateral.

The project returns depend on effort selected by E. This is noncontractible (i.e., effort covenants cannot be written in the contract) either because of difficulties L faces in monitoring E’s effort, or because this cannot be verified by third party contract enforcers. If E does not default, his net wealth is $R - l(1+r) + w$ (where R denotes the project return), while if he defaults it is 0. Hence E’s stake in avoiding default equals the difference in net

wealth between the two states: $S = [R - l(1 + r) + w]$. The higher this personal stake, the greater will be E's inclination to avoid default. Hence the default probability d will be a decreasing function of $R - l(1 + r) + w$.

The lender's return of L will be $l(1 + r)$ in the event of no default, and $R + w$ in the event of default. For simplicity assume that the project has only one of two possible outcomes: a return of R in case of success, and 0 in case of failure. The total expected return to L will then equal

$$d(S)w + [1 - d(S)]l(1 + r) = d(S)[w - l(1 + r)] + l(1 + r) \quad (1)$$

which in turn should equal the lender's cost of capital in a competitive credit market.

As expected, the lender's expected return is decreasing in the default probability (since by hypothesis w is smaller than l). Equation (1) immediately yields the following implications:

Wealth Effects: Higher wealth w allows E to post more collateral, increasing his personal stake S in avoiding default, besides reducing the lender's exposure in the event of default. Hence wealthier borrowers tend to work harder to avoid default, unless there are significant wealth effects on the demand for leisure. Absent this latter complication, wealthier borrowers work harder and so incur lower default risk, thus generating higher returns for lenders *ceteris paribus*. In the presence of wealth effects on leisure preferences, it can be shown (see Mookherjee (1997b)) that the effect of increasing wealth is to increase effort or leave it unchanged over an initial range of low wealth levels (where the limited liability constraint binds), and then decrease it over the rest of the wealth range (where the liability constraint does not bind). The effects on the lender's expected profits are analogous. The result stating that over some range lenders prefer to deal with wealthier borrowers, is particularly significant, since it implies a complementarity between internal finance and access to external finance. It highlights the role of limited liability constraints, since in their absence lenders would generally prefer to transact with poorer agents (as shown by Newman (1995) and Thiele and Wambach (1998)).

Figure 2: EFFECTS OF VARYING INTEREST RATE

Figure 3: UTILITY POSSIBILITY FRONTIER

Interest Rate Effects: Higher interest rates now have two effects: (i) transfer income from E to L in the event of solvency, and (ii) an incentive effect, since they increase the debt burden on E, reducing the latter's stake S in solvency, thereby inducing default risk

to increase. The first of these is a pure transfer, while the second involves deadweight loss. E does not internalize the adverse effect on L of a higher default risk, the source of the moral hazard problem, and thus tends to exert too little effort. Higher interest rates tend to increase this deadweight loss. The effects of higher interest rates on the *ex ante* utilities of the two parties is sketched in Figure 2. The resulting utility possibility frontier is sketched in Figure 3. Of particular interest is the phenomenon that the lender's profits do not increase monotonically in the interest rate: beyond some ceiling rate r^1 , further increases in the interest rate lower lender's profits by causing the second effect on default risk to outweigh the first transfer effect.

Informational Rents: Owing to this interest rate ceiling, all borrowers with wealth smaller than some threshold (w^1 in Figure 3) will earn a surplus over and above their outside option of autarky. The purpose of this surplus is to provide the borrower with sufficient incentive to avoid default which makes both parties worse off. It cannot be extracted by a monopolistic lender because of the limit on the borrower's liability. These rents have significant welfare implications, since they represent an externality between L and E. L in general does not incorporate the benign effect of lower interest rates on E's rents, and thus tends to set them too high from an efficiency standpoint. In turn they have the following consequences.

Role of Credit Market Structure and Property Rights Allocation: Suppose that the credit market is monopolistic and the borrower is poor: then the interest ceiling rate of r^1 will be charged, which maximizes the expected profit of the lender. If the market is competitive, the expected utility of the borrower will be maximized instead, where the interest rate is substantially lower. Hence increased competition has the effect of lowering the deadweight loss associated with a high default risk. More generally, any parametric change which expands the bargaining power of the borrower (as defined by the implicit welfare weight assigned to E, a rightward shift along the utility possibility frontier) causes an efficiency improvement. This may include a reallocation of property rights over relevant productive assets from L to E, or an expansion in E's outside options (see Mookherjee (1997a)) — either of which could result from redistributive state policies.

Nonlinear Interest Rates and Credit Rationing: Expansion in loan size to finance higher levels of consumption of E at the outset of the project will increase the debt burden $l(1+r)$, without expanding the returns from the project. This will reduce E's stake in success, causing default risk to increase. This may outweigh the effect of a larger scale of borrowing to imply that the lender is worse off. The larger loan must therefore be accompanied by a different interest rate and/or level of collateral, which is the phenomenon of nonlinear interest rates. Increases in the interest rate can however make matters worse, by raising debt burdens even further. Hence, the lender may be unwilling to lend beyond some level of loan size at *any* interest rate, which is the phenomenon of credit rationing.

Role of Long-term Exclusive Relationships, Interlinked Contracts, and Social Networks: When the lender and borrower enter a long-lived relationship, it expands the opportunity for the lender to relax limits on the borrower's current liability by extracting repayment in future successful periods (by the institution of debt), or by the threat of terminating the supply of credit (Dutta, Ray and Sengupta (1989)). In general, all other transactions the borrower enters into, either with other creditors or on markets for related goods and factors, will have spillover effects on effort incentives with respect to the project being financed by any given lender. This creates an externality between creditors and other agents trading with E, motivating exclusive dealing clauses, whereby the same party (L) transacts on all fronts simultaneously with E (Braverman and Stiglitz (1982), Kahn and Mookherjee (1998)). Moreover, it expands the opportunity to relax the limits on E's liability with L in default states, by taxing other transactions in the event of default. A similar reason underlies the role of lending within social networks, where punishments can be imposed for loan defaults in other spheres of social interaction, and third-party community-based sanctions can be brought to bear on defaulters owing to the rapid flow of information within the community.

2.1.2 Empirical Evidence

As described above, the presence of positive wealth effects on credit access or on aspects of contractual structure facilitating high effort incentives. Two recent studies are noteworthy in this respect.

Laffont and Matoussi (1995) directly test such a model in the context of sharecropping contracts in Tunisia. Specifically, they examine: (i) the effect of working capital and wealth of tenants on the likelihood of contract choice involving fixed rent or a high share for the tenant (which induce high effort incentives), and subsequently (ii) the effect of contractual choice on total factor productivity. Econometric problems of endogeneity are handled by using age, wealth, family size, and owned land as instruments. They find significant positive effects of the working capital and wealth of tenants on the likelihood that a high-powered contract will be chosen, thus indicating the relevance of limited liability constraints. Moreover, contractual choice has a significant effect on farm productivity: moving from a sharecropping contract to a pure rental contract expands output by 50%.

Akerberg and Botticini (1998) examine the effects of crop risk and tenant wealth in 15th century agriculture in the Italian province of Tuscany, while controlling for endogeneity created by matching between landlords and agents of heterogeneous risk-aversion and wealth characteristics. Irrespective of whether or not these endogenous matching incentives are controlled for, they invariably find that tenant wealth has a positive and statistically significant effect on the likelihood of use of a fixed rent contract, thus suggesting a positive correlation between wealth and effort incentives (if income effects in the demand for leisure can be ignored).

These micro-econometric analyses of capital market imperfections and the importance of ‘internal finance’ vis-a-vis external finance, are complemented by numerous historical accounts of the role of ‘primitive accumulation’ as a precondition for successful industrialization. This refers to the accumulation of large investible surpluses — either in agriculture, trade or other means — by an entrepreneurial class. Early industrialization is rarely based

on externally borrowed funds. Two relatively recent examples (to be reviewed further in later sections of this paper) are the Western Indian sugar cooperatives in the 1940s and 1950s, and the Southern Indian garment export industry of Tirupur since the 1970s. In both cases the origins of the industry were represented by the adoption of lucrative cash crops at the turn of the 20th century — sugarcane and cotton respectively, both stimulated by the advent of the railways that created national markets, and by large scale expansion of irrigation by the government — by certain communities that were to later transform themselves into dynamic industrial entrepreneurs (e.g, see Chithelen (1985) and Banerjee-Munshi (1999)). Given weak formal financial markets, and problems of monitoring and enforcing loans to distant communities, these capital funds were rarely invested outside the local community. Given the scarcity of land, the only opportunity for investing them profitably required the starting of factories that would enter downstream production and market the corresponding finished product. As Banerjee and Munshi document, the Gounder community invested more in capital equipment than members of other ‘outside’ communities, and they argue this eventually played a significant role in their steeper export growth rates.

Capital requirements for setting up factories, relative to the level of surplus available to individual farmers, also had a significant impact on the current organizational forms of these industries. In the case of the Western Indian sugar industry, the majority of the sugarcane farmers were peasant cultivators that did not own the capital necessary to integrate forward to the crushing and processing of sugar, giving rise to the need to set up cooperative factories, often with extensive startup capital contributions from the government. In contrast, the sugar industry in the central Indian state of Uttar Pradesh was started by a trader class that acted traditionally as sugarcane marketing agents, that were substantially wealthier, thus possessing sufficient individual capital to start private sugar factories (Amin (1984)). The differences in the level of investible surpluses of the entrepreneurs involved is perhaps the single most important explanation of why the sugar industry is primarily composed of cooperatives in Western India and private factories in central India. In contrast to sugar, the garment industry involves many successive stages of manufacture, each of which is characterized by lower economies of scale. This made it affordable for peasant farmers to invest in equipment necessary to enter a single stage of the garment produc-

tion process, while entering into an extensive community-based network of subcontracting that could collectively deliver the finished product. Hence the Tirupur garment industry managed to evolve without external state assistance, relying instead on a semi-cooperative network structure. As individual firms grow and accumulate greater internal funds, there is an expressed interest in integrating forwards and backwards (Cawthorne (1995, p. 48)).

2.2 Investment Implications of Credit Market Imperfections

Capital market imperfections and related wealth effects have profound implications for investment incentives. For instance, if the cost of capital is higher for poorer entrepreneurs, it reduces the rate of return they can earn from their investments, relative to wealthier entrepreneurs. This implies that neoclassical forces toward wealth convergence are mitigated. The poor may not be able to invest and grow their way out of poverty, whether in physical or human capital, preventing catch-up with the wealthy. Dynamic models probing such issues in contexts under the assumption of constant savings rates, have been studied by Banerjee and Newman (1993), Galor and Zeira (1993), Aghion and Bolton (1997) and Piketty (1997). In most of these models, there are multiple steady states, with varying levels of long run inequality and per capita income. High inequality tends to be associated with low levels of per capita income, owing to the investment-inhibiting effect of low levels of wealth.

A common ingredient in these models has been the assumption of an exogenously fixed savings rate, and indivisibilities in investment.³ The intergenerational wealth dynamic in the Galor-Zeira model, for instance, is depicted in Figure 4. There are two stable steady states, one associated with low wealth (w_1) and no investment in education, and the other with high wealth (w_3) and investment. Families with intermediate wealth levels do invest in education but need to borrow to finance these investments. Since the borrowing rate of interest exceeds the lending rate of interest, reflecting the capital market imperfection, the intergenerational wealth dynamic has a steeper slope for these intermediate wealth

³The Piketty model allows for continuous investment, however, but does rely on a constant savings rate.

Figure 4: WEALTH DYNAMICS IN GALOR-ZEIRA MODEL

levels. This creates the nonconvexity that allows for multiple steady states, including an intermediate (unstable) one at w_2 . Families that start historically with a wealth below w_2 will end up at the low-wealth no-investment steady state w_1 , while those that start above w_2 will converge to the high-wealth education-acquiring steady state w_3 . Note that for a family with wealth slightly below the unstable steady state level w_2 , a small increase in its current saving would tip its succeeding generations permanently into the class of the wealthy. Hence a small degree of altruism or foresight would tip such families out of the low level trap. The assumption of a constant savings rate prevents current generations from making small sacrifices even though their long-run effects might be highly significant. In that sense the result of the low level investment trap seems to be driven by an assumption of excessively myopic behavior.

In addition, the assumption of indivisible investments implies that it is not possible for a poor household to slowly build up its stock of human capital over time. This tends also to inhibit the process of investment somewhat artificially. In the Galor-Zeira model, for instance, a poor household must decide whether or not to educate its children, and if it does not, its children are in exactly the same position as their parents were (in a steady

state). Smooth patterns of upward mobility in skill acquisition are simply ruled out by assumption. The role of this assumption is illustrated vividly in the model of Loury (1981), where there is long run convergence to a unique ergodic wealth distribution, as a result of a continuous investment technology, despite the complete absence of any capital markets.

These considerations suggest the need for models which examine the question of long run convergence, with endogenous savings rates, and a continuous investment technology. Ljungqvist (1993) allows savings rates to be endogenously determined in an intergenerational model of human capital acquisition, while continuing to assume that the educational investment technology is discrete (i.e., either a family educates its children, or they do not). He shows that there will typically be multiple steady states where inequality and per capita income are negatively correlated. In this sense the result of preceding models that assumed a constant savings rates continues to be vindicated. Mookherjee and Ray (1998) extend the same model to the context which additionally admits continuous upgrading of skills. They find that there will typically be a unique steady state marked by long run inequality and complete absence of any mobility. In other words, the result concerning persistence of inequality continues to obtain, but not the multiplicity of steady states. Both these papers follow Loury (1981) in assuming that there are no capital markets present at all.

Ghatak, Morelli and Sjostrom (1998) and Mookherjee and Ray (1999) in contrast examine similar issues in a context where an (endogenously) imperfect capital market does exist, where the imperfection stems from borrower moral hazard with limited liability. In particular, the latter paper studies a context with an endogenous and continuous investment rate, so we give a brief account of the results of this model.

In the simple static context where there are no wealth effects on the demand for leisure, and borrowers are risk-neutral, note that both lenders and borrowers are better off when the borrower becomes wealthier — higher wealth permits larger collateral, inducing credible commitment to higher effort levels, which reduce the default risk. Hence wealth accumulation by the borrower would be Pareto improving. This would require the borrower to aggressively plough back profits into the retained earnings of the enterprise, abstaining

from paying them out in dividends, until a point that the entrepreneur would have no further need for external borrowing. This would lower the cost of capital in the long run, permitting the firm to grow faster. Under what circumstances would the credit market encourage such patterns of accumulation?

Let w_t denote the level of wealth/internal finance/retained earnings of the entrepreneur at date t , now to be treated as an endogenous variable. Assume that the project scale is not changing over time, so the input requirement is fixed at f in every period. If w_t is less than l the entrepreneur needs to borrow at t , not otherwise.

We assume that all borrowing is intraperiod, i.e., provided for meeting working capital needs; the possibility of more long term interperiod debt does not exist. This is the simplest dynamic extension of the single period model. In particular, the market offers only such short-term contracts, which can be constantly renegotiated. We shall seek to identify Markov equilibria in this setting, where credit contracts at any date are conditioned only on the enterprise's current wealth level; all other past history is ignored.

Suppose that the entrepreneur does succeed in getting a loan at interest rate r_t , and proceeds with production. As in the one period setting, the returns are stochastic: if successful it returns R and repays the loan, while if it is a failure it returns nothing and the entrepreneur defaults on his loan. Hence the wealth level following production at date t is given by $I_t = R - l(1 + r_t) + w_t$ if solvent, and 0 otherwise.

In this setting, we have to introduce an accumulation decision at the end of every date. Specifically, if the enterprise has attained a total wealth of I_t following production at t , how much of this should be paid out in the form of dividends d_t that will be consumed at t ? Assume an exogenously given interperiod lending rate of interest i , which also equals the discount rate of future utility (thus ensuring that in the absence of credit market imperfections and a smooth income flow the entrepreneur would have no intrinsic desire to save or dissave). Then the retained earning at the end of period t will be $I_t - d_t$, and at the beginning of period $t + 1$ will be $w_{t+1} = (1 + i)(I_t - d_t)$.

What incentives does the entrepreneur have in such a world to plough back earnings into the enterprise? It would provide a greater stock of internal finance in future periods, reducing the need to rely on external borrowing. This would induce greater efficiency in future periods, by reducing the enterprise's future debt burdens and thus motivating the entrepreneur to exert greater effort. The answer depends on the extent to which the returns from this extra efficiency accrue to the borrower.

Suppose first that the credit market is competitive, so credit contracts maximize the present value utility of the entrepreneur, subject to a period-by-period breakeven constraint for lenders. Then the entire returns accrue to the borrower. Absent any intrinsic desire to smooth consumption, the entrepreneur will want to plough back *all* profits, abstaining from paying out any dividends at all, so the enterprise will accumulate internal finance at the fastest possible rate. The result is that the retained earnings of the enterprise follow an ergodic process, converging to a long run distribution independent of initial conditions. In particular, entrepreneurs that start poor might (with positive probability) become rich in due time: there are no poverty traps. On the other hand, cash-rich enterprises without any need for external borrowing may suffer a string of failures which eliminates their cash reserves, but they will subsequently tend to rebuild their reserves at the fastest possible rate. The long run ergodic property and the absence of any poverty traps happens to survive even when the entrepreneur has concave utility, and irrespective of whether or not the time horizon is finite or infinite.

In contrast, consider the effects of a monopolistic credit market, i.e., credit contracts maximize the welfare of the lender, subject to a participation constraint for the borrower. Then part of the benefits of accumulated internal finance of a credit-constrained borrower accrue to the lender. Indeed whenever the borrower starts with low initial wealth, *all* the benefits accrue to the lender. This follows from the informational rent property noted in the one period model. To illustrate the argument in more detail, consider a two period setting, where both borrower and lender are risk-neutral. Figure 5 describes the expected date 2 utilities of the two parties as a function of the borrowers wealth position w_2 at the beginning of that date (this is simply the result of applying the optimal static contract

Figure 5: EFFECT OF VARYING BORROWER WEALTH

described above). When it falls below the threshold w^1 , the borrower applies the ceiling interest rate r^1 , allowing the borrower to earn an informational rent. Small increases in his wealth position w_2 will still allow the borrower to earn such a rent, so the lender will continue to apply the same interest rate. The borrower's effort and utility is then left unaffected, while the lender benefits by being able to secure a larger collateral in the event of a default. The benefits of higher wealth of the borrower thus accrue entirely to the lender; the monopoly power of the lender effectively operates as a 100% tax on marginal wealth accumulation by the entrepreneur. At higher wealth levels (i.e., above w^1) the borrower's net income increases exactly in step with his initial wealth, as no further rents accrue (this is just a consequence of the fact that the utility possibility frontier in Figure 2 is downward sloping after the point corresponding to the ceiling interest rate r^1). Hence the entrepreneur is exactly indifferent about increases in future wealth beyond w^2 achieved as a result of cutting back on dividend payouts. It follows that irrespective of the starting wealth w_1 at date 1, it is optimal for the entrepreneur to consume everything at date 1, so as to start date 2 with zero reserves.

The argument is straightforward to extend when there are an arbitrary finite number

of periods: the same argument will apply to the last date, so the entrepreneur will consume everything at the penultimate date. Then the penultimate date acts effectively as the last date and the whole argument applies again to the date immediately preceding it. And so on: it follows that the unique outcome will be for the entrepreneur to *never* plough back any earnings into the business. In other words, there is a poverty trap: zero wealth state is an absorbing state. It can be shown that the argument extends also in suitable ways when the horizon is infinite and the entrepreneur has concave utility. The dynamic inefficiency resembles somewhat the inability of a durable goods monopolist to capture any consumer surplus, known as the Coase conjecture. It owes to the inability of the monopoly lender to commit to not appropriating increased cash reserves of the enterprise in future periods.

What does this imply for the role of interventionist credit policy? Note that the government will have to effectively compete with lenders that exercise monopoly power. To the extent that this monopoly power of private lenders stems from their superior monitoring and enforcement abilities, the government will be hard placed to exert any competitive pressure. Indeed, government credit programs in developing countries have frequently encountered severe problems with the enforcement of loan repayments, so credit programs have frequently been converted to income transfer programs. The model suggests that such transfers will be consumed rather than invested, so the interventions would be ineffective in getting poor entrepreneurs off the noninvestment trap. Effective interventions will then require either substantial reforms in monitoring and enforcement by public sector banks or nonprofit institutions to enable them to compete with existing private lenders, or measures to encourage the development of financial market institutions that expand external financing options available to small entrepreneur.

An added complication in the context of a competitive credit market is pointed out by Ghatak, Morelli and Sjostrom (199?): policy measures to alleviate credit constraints will have the perverse effect of discouraging savings and effort in preceding periods, since entrepreneurs will gain less from relieving their own credit constraints.

3 Joint Ownership and Two-Sided Moral Hazard

Lin and Png (1998) quote a leading Taiwanese businessman:

“One of my friends has two factories in Mainland China. Both are exactly identical, except that one was a joint venture with the local authorities. One time, there was a shortage of electric power. His wholly owned factory had to shut down. But the joint venture factory had lot of electricity.”

The problem is that there are two inputs supplied by distinct parties: capital and some material input. Capital inputs have to be provided first in a lumpsum and are sunk, while material inputs will be supplied on an ongoing basis later. If long term contracts cannot be signed or enforced, whence the terms on which the material input will be supplied later are agreed on in advance, the material input supplier can ‘hold up’ the capital investor *ex post* by demanding a higher price or delivering low quality supplies. Anticipating this, the investor is unwilling to invest much upfront, and the result is an underinvestment trap. This problem has been the subject of numerous theoretical contributions, most frequently associated with the work of Williamson (1975, 1985), Crawford, Klein and Alchian (1978), Grossman and Hart (1986) and Hart and Moore (1990). These authors argue that one solution is to change the structure of property rights, i.e., either the input supplier owns the firm and contracts for capital with an independent investor, or the input supplier and capital investor jointly own the firm’s assets. Alternative solutions may involve the use of contracts which safeguard the interests of foreign investors against *ex post* expropriation, or the development of reputational forces that prevent such forms of short term opportunism. Of these two, the first is unlikely to be effective given the weak legal protection available to foreign investors in China. Reputational concerns require a long term horizon for the Chinese state authorities, and resistance to short term populist or fiscal pressures to squeeze foreign investors. In this light, altered property right structures may be a more reliable way to safeguard the interests of foreign investors, even though this may require the sacrifice of some of the rents on their part to their local hosts.

To the extent that the role of the investor is only to provide finance, exclusive ownership by the local state will indeed solve the problem. However, in many instances, the investor is also an entrepreneur, who manages the enterprise and markets the product in export markets. Such skills may be beyond the scope of state officials. These managerial and marketing activities are subject to moral hazard, just as much as the input supplier is prone to *ex post* opportunism. For instance the entrepreneur has to be encouraged to work hard in setting up the enterprise, in keeping costs under control, and in aggressively exploring market opportunities. With such double-sided moral hazard, exclusive ownership by the material input supplier may result in inefficient management of the enterprise. The appropriate solution may then involve joint ownership, which makes a compromise between the two moral hazard problems.

Lin and Png explore the role of this hypothesis in explaining the pattern of direct foreign investment by Taiwanese investors in mainland China. They argue that the scope for *ex post* opportunism would be attenuated in the presence of kinship ties between the two parties concerned, as people would be less inclined to chisel on those with whom they share common kinship or family ties. To the extent that the Chinese state authority officials belonged to the same kin as the Taiwanese investors, the risk of unreliable electricity supply to wholly foreign owned enterprises would be less acute. Hence in such locations, the investments would tend to take the form of wholly foreign owned enterprises rather than joint ventures where the local state was a co-partner.

They attempt to test this proposition with data on direct investments by Taiwanese nationals in China between 1987-91. There is a close kinship identity of these investors with Fujian province on the southeastern coast of China: 80% of the Taiwanese population are of Chinese origin, as Fujian province is geographically the closest to the offshore island. Geographical proximity, however, does not translate into lower transport costs and ease of monitoring, since during this period the Taiwanese government regulations officially prohibited Taiwanese investors from directly investing in China. Consequently all the investments had to flow through Hong Kong, which was further south along the coastline, and the relevant measure of economic distance was location relative to Hong Kong. From this

standpoint, investments in Fujian province were actually further away from Taiwan than other provinces such as Guangdong. This enables Lin and Png to distinguish between the holdup hypothesis and the alternative (simpler) explanation in terms of lower monitoring and transport costs.

The data set does seem to provide positive support to the holdup hypothesis: Taiwanese investments in Fujian province were 20-22% more likely to be wholly foreign owned, controlling for a host of other determinants of ownership form, such as distance from Hong Kong, scale of the enterprise (larger scales necessitating larger and riskier investments, thus encouraging joint ownership), and Chinese government regulations for foreign enterprises (such as Special Economic Zones which provide superior infrastructure, where the scope for holdup is smaller). In addition, distance from Taiwan had no effect on investments outside Fujian province, after controlling for distance from Hong Kong, so that anticipated relaxation of the Taiwanese curb on investments in China cannot explain their result. They also show that their result survives even controlling for the benefits of a common language, based on survey responses of the Taiwanese investors.

The implications of the holdup hypothesis are that such problems inhibit investment. To the extent that the Chinese industrial growth pattern has relied substantially on investments and export marketing skills of offshore investor-entrepreneurs, joint ventures were one solution to the holdup problem, but a costly one for investors as it required them to sacrifice some of the rents from these enterprises. Kinship ties with the locations of their investments moderated the holdup problem, which would presumably have encouraged the flow of such investments. Of course direct evidence on the effect of kinship ties on the volume and rates of return on direct foreign investments (rather than their organizational form) is not available: such studies would be welcome in future research. Countries less fortunate in terms of the location of external investors with common kinship ties — as most other developing countries are — would have to seek alternative measures to moderate the holdup problem, such as strengthening legal protection for foreign investors, and committing to better infrastructural facilities.

3.1 Two-Sided Moral Hazard with Limited Liability: Agro-Processing Cooperatives

Agroprocessing industries play an important role in industrialization for obvious reasons; they usually represent the first stage of transformation from an agrarian-based economy to an industrial society, and capitalize on historical comparative advantage in specific cash crops. Examples are cotton, rubber, dairy, sugar, tobacco, coffee, oilseeds and related oil products, canned food (e.g., shrimp) and fruits (e.g., pineapple, banana). The processing of most of these has high value-added, and requires substantial investment in machinery. Some of them are less perishable than others and are easier to transport (e.g., cotton or oilseeds), and so can be processed and turned into a finished product (e.g., textiles or oil) in factories located at large distances from the crop source. Other crops are perishable and difficult to transport, or need to be processed very quickly following the harvesting, such as rubber, sugarcane, dairy and fruits. These crops therefore need to be processed near the crop source. In the latter case, a comparative advantage in the crop translates naturally to a comparative advantage in the corresponding processing industry.

Processing naturally involves two-sided moral hazard, with the two parties being (i) the supplier of capital, management and marketing expertise; and (ii) the supplier of the crop. When the two parties are distinct, and the crop has to be processed quickly following harvesting, it creates an opportunity for hold-up of the farmer by the processing factory. This is the result of economies of scale in processing: each factory is typically large and therefore located at some distance from each other. Efficiency dictates that each factory transact only with growers in the area it is surrounded by, owing to the need to keep the time lag between harvesting and processing down. This creates an *ex post* bilateral monopoly situation where each party can get held up by the other. Ideally, the harvesting and crushing operations need to be scheduled harmoniously, a logistically daunting task when the crop is cultivated by numerous peasants. Such *ex ante* schedules can get upset *ex post* as a result of opportunism by either party. In years of crop shortage, farmers can hold-up the factory by threatening to send their crop elsewhere if the factory does not offer a higher price than previously promised. The factory would not be able to fill the resulting gap in its processing

schedule at short notice, resulting in its equipment not being efficiently utilized. On the other hand, when there is a large harvest, the factory owner can exercise local monopsony power and offer a lower price for the crop than previously promised. The opportunity for the farmer to take his crop elsewhere at that point is limited, since other factories have already scheduled their processing operations at full capacity. When such forms of *ex post* opportunism are rampant, the result is frequently low supplies of the crop (especially if the factory owner cannot commit to pay a remunerative price to the growers), poor quality of the crop (resulting from large time-lags between harvesting and crushing) or low levels of capacity utilization (when the scheduling between harvesting and growers is not efficiently organized, or when growers hold up the factory).

As the discussion of the preceding section suggests, one solution to this is for the crop cultivation and processing operations to be jointly owned, which would eliminate *ex post* holdup and allow efficient scheduling of harvesting and processing. Correspondingly there are three possible distinct organizational forms in these industries:

Plantation systems represent such a form of common ownership, where the capital suppliers integrate backward into crop cultivation. The problem with this is that it typically involves a larger investment (since the factory has to purchase the land required), a larger range of operations for factory managers with possible diseconomies of scope, and inefficiency in cultivation (owing to a replacement of family labor by wage labor). Allowing factory owners to integrate backwards frequently result in displacement of numerous peasant cultivators with resulting increase in local inequality and political tension. They run counter to professed social objectives of many democracies, which frequently impose ceilings on landownership.

Farmer cooperatives represent an alternative solution, which tend to be more egalitarian, or at least appear to be so. Here peasant cultivators integrate forward to common ownership of the factory. This retains the advantage of cultivation utilizing family labor, while the principal disadvantage is that management of the factory is subject to greater inefficiency as a result of its diffused ownership. These problems include

possible heterogeneity and interest conflict among diverse groups of members, and problems of coordination and monitoring of factory management. Starting cooperatives also involves overcoming problems of collective action, besides possible capital shortages. For this reason, the government has frequently taken an active role in the sponsorship and subsidization of such cooperatives. In many developing countries (e.g., in Africa) the government has also played a role in the management of these cooperatives.

Privately Owned Factories involving separate ownership of factory and lands represent a half-way house, which is subject to the two-sided moral hazard problems described above. But it does limit the need for larger investments and scale of operations that the plantation system is associated with, and allows both processing operations and cultivation to be efficiently organized because each is more closely held, thus limiting distance between owners and managers/workers.

This discussion suggests that each organizational form has corresponding strengths and weaknesses. So it is unsurprising that diverse ownership structures are observed to coexist in different countries or locations within the same industry.

Banerjee, Mookherjee, Munshi and Ray (1997) (hereafter referred to as BMMR) have investigated the performance of sugar cooperatives in the Western Indian state of Maharashtra. India is the leading sugar producing country in the world, and Maharashtra is the leading producer amongst all Indian states. Over 90% of Maharashtra's sugar is produced by cooperatives, which are among the most competitive and dynamic sugar producing operations in the world. Sugarcane yields as well as recovered sucrose output per hectare were higher in Maharashtra than any other sugar producing region in the world (i.e., the rest of India, Australia, Brazil, Cuba, Europe, Mexico, South Africa or the USA) in the 1980s as well as early 1990s (Sawhney (1997)).

While analyses of the performance of cooperatives *vis-a-vis* privately owned factories are still awaited (I have research currently in progress with Sanghamitra Das on this topic within the Indian sugar industry), BMMR concentrate on the determinants of performance of

cooperatives within different parts of Maharashtra. The factories in the state can be divided into two distinct regions, the Eastern and Western parts, where sugarcane cultivation in the former has a distinct advantage with respect to soil, topography and irrigation. Yet, surprisingly, the West dominates with respect to both levels and growth rates of capacity of the sugar factories. These cannot be explained by differences in either crop quality or crushing efficiency: the recovery rates (of sugar from sugarcane supplies) in the two regions do not differ much, nor do they exhibit any time trend.

One explanation for the smaller capacity levels and growth rates in the East arises from problems of collective decision making arising from greater heterogeneity among its members, specifically between small and large farmers. Large farmers typically start the cooperative and manage it, while the small farmers collectively supply the lions share of the sugarcane. The large farmers have greater control over profits of the enterprise, being able to divert it for their private uses in various forms, and thus seek to depress the price paid for sugarcane delivered by the small farmers. Being aware of this proclivity, state regulations forbid discriminatory prices to be paid for the produce of different members. So the large farmers must pay themselves a lower price if they attempt to lower the price for the small farmers. From their point of view, therefore, the optimal price trades off the higher profits recovered from the small farmers, against the lower revenues they recover from their own harvest. Clearly, the greater the fraction of land owned by small farmers, the lower the optimal price from the point of view of the large farmers, creating a larger distortion relative to the first-best cane price. On the other hand, the small farmers always prefer a high cane price to a lower one.

This interest conflict is resolved according to the relative numerical strengths of the two groups, given the role of majority rule in shareholder meetings, and in elections to the controlling Board. If the small farmers become more numerous, their interests are better represented. BMMR show that the result is an under-pricing distortion which is U-shaped with respect to the ratio of small to large farmers, as depicted in Figure 6. For example, if this ratio were zero, the cooperative would be perfectly homogenous, consisting only of large farmers. Since there would be no small farmers to exploit, the selected price would involve

Figure 6: CANE UNDERPRICING AND COOP COMPOSITION

Figure 7: CAPACITY LEVELS AND COOP COMPOSITION

no distortion. A small increase in the number of small farmers would make it optimal for the large farmers to underprice to some degree, and they would continue to be able to set the price since they still retain control rights. However, as the number of small farmers

grew beyond some level, the large farmers would gradually lose control, causing the selected price to go up. In the limit as the small farmers comprise the entire membership, all interest conflicts would be resolved, and the distortion would again disappear.

The central prediction of the model concerns the role of local landholding patterns on the degree of interest conflict within the cooperative, and the resulting distortions in the pricing and supply of sugarcane. A U-shaped pricing distortion will imply a corresponding U-shaped supply pattern. Since capacity tends to be matched with the expected cane supply that will need to be crushed, it will also tend to be U-shaped, as depicted in Figure 7.

These predictions are upheld in an examination of the data concerning pricing and capacity of individual cooperatives. The econometric analysis controls for potential endogeneity of participation rates and landownership distributions, besides using other properties of the data to discriminate between heterogeneity and alternative hypotheses. As it turns out the Eastern part of the state is characterized by landholding distributions where large farmers occupy a larger role, near the bottom of the U. Small growers are more numerous in the Western part, pushing them closer to the right end of the U. This explains the larger capacity levels in the East. The results thus suggest the role of unequal landholding patterns in raising wasteful conflicts of interest between large and small farmers in the cooperative. The geographic comparative advantage of the Eastern part is outweighed by its comparative disadvantage with respect to institutional attributes.

The analysis also explains the patterns of capacity dynamics. Over time landholdings have tended to become more fragmented over time, in both the Eastern and Western parts of the state. This reduced interest conflicts and induced capacity growth within both regions. But since the slope of the U is much flatter near the bottom than on its right arm, the dynamism imparted was stronger in the West, resulting in a higher growth rate.

4 Reputational Mechanisms

4.1 Why Quality and Reliability are More Problematic in LDCs

One of the common problems of industrial enterprises is to assure customers of a reliable, quality product. Stories of adulterated foodstuffs, equipment breakdowns, and delays in delivery are commonplace, and constitute a significant drawback in their competitiveness with comparable products from developed countries. The most common explanation of these problems of low quality is that customers in developing countries are poorer and thus less willing to pay for high quality. In many contexts this explanation is not persuasive, e.g., if customers prefer to purchase low quality milk, why should milk suppliers dilute the milk when the customers can do it themselves? Why don't customers of comparable wealth in richer countries get served in a similar fashion? If delivery is to be delayed why should the customer not be warned in advance, instead of being kept in limbo? One suspects that some broader, systemic problems are involved. For instance, reputational considerations play a significant role in maintaining quality incentives for experience goods with significant incidence of repeat dealings. A customer served adulterated milk by a given supplier is unlikely to continue patronising that supplier in the future: for most customers the incentive problem cannot realistically be solved by integrating backward into the production of milk (though there are some customers who are known to do so for precisely this reason!).

This raises the question: why does such reputational forces work less effectively for producers of experience goods in developing countries? Esfahani (1991) provides one explanation, in terms of greater variance in the costs of providing high quality products. He develops a reputational model on the lines of earlier work by Klein-Leffler (1981) and Shapiro (1983), where a sector providing high quality products coexists with a competitive fringe providing low quality products. For reputation to work as an effective carrot, the supplier of high quality needs to be rewarded with a price premium above the cost of supplying high quality, which would be foregone in the future following an opportunistic deviation to

low quality today. Such premiums are consistent with competition which eliminates excess profits for firms: a firm newly entering the market offers an introductory bonus, pricing its product below cost. This initial loss is recouped by premiums earned later. Any deviation to low quality supply will ensure loss of clientele in the future, who will switch to a newly entering firm, or to low quality service produced by a competitive fringe.

In this context, suppose that the marginal cost of producing a high quality product is subject to greater uncertainty. The cost is observed by the producer before making the decision concerning the quality to be supplied in any given period, but cannot be observed by customers. Then customers cannot graduate the future punishment for provision of low quality to the current cost of high quality, e.g., make exceptions owing to ‘extenuating circumstances’. In other words, the *same* punishment will be meted out by customers, irrespective of the current cost shock. So the existence of a reputational equilibrium relies on reining in the incentive to skimp on quality in the *worst* possible cost state. Widening the range of possible cost shocks however increases the current temptation to deviate. This can be controlled only if the future quality premium were to increase. This tends to cause customers to switch to low quality substitutes, and the reputational equilibrium threatens to fall apart.

Why might costs of producing high quality goods be more variable in developing countries? Though this appears a realistic hypothesis, it is not itself explained by the Esfahani model. But it does suggest an interesting extension, where production has a circular Leontief-like structure, with commodities used to produce commodities. Then if the producer of any given commodity experiences greater variability in the supply of inputs, it makes it more difficult for him to sustain a steady reputation for high quality, thus rendering the supply of his own commodity more variable to his customers. On the other hand reliable supplies of his own inputs would induce high quality provision of his own good. Hence there is the possibility of *multiple equilibria* in reliability. Moreover, to the extent that certain key infrastructural inputs are delivered by the state, such as transport, power or communications, improvements in the reliability of such services may well have impressive multiplier effects on the reliability of private sector products.

Esfahani points out a number of other implications of his model:

greater capital scarcity in developing countries can induce higher discount rates, reduce the present value of future quality premiums, and thereby undermine incentives to maintain a reputation for quality;

large firms have an advantage over small firms with respect to building a reputation: they can vertically integrate to increase input supply assurance, and have lower discount rates owing to larger reserves of internal finance, and lower turnover rates

increased competition on the product market can exacerbate quality problems: oligopolistic structures help protect quality premiums from the threat of competition by new entrants, at the cost of an increase in prices

4.2 Reputational Externalities

An interesting extension of the simple reputation model suggests the existence of pervasive externalities across different sellers of a given product. These are likely to arise particularly when there are multiple sellers and agent turnover prevents the formation of long-term relationships between particular seller-buyer pairs. Tirole (1996) presents a model of *collective reputation*, where buyers of a product cannot distinguish different sellers from one another, and form beliefs about a representative seller they might encounter in the future based on their past dealings with different sellers. Then an opportunistic deviation by one seller to low quality would cause the buyer to revise her beliefs about the quality provided by future sellers in a pessimistic direction. This would reduce their willingness to pay for the product of future sellers, thus destroying the premiums for quality necessary to preserve reputational incentives. In this context, pessimistic historical beliefs held by importers concerning the quality of product sold by exporters of a given developing country will prevent an export take-off despite favorable changes in export competitiveness.

Another variation is to allow for heterogeneity in the set of agents subject to moral

hazard. For instance, suppose that in the seller population, some are ‘fly-by-night’ operators who are entirely myopic, and would therefore not be swayed at all by the prospect of losing future business. As Ghosh and Ray (1996,1998) and Kranton (1996) have argued, nonmyopic sellers would take time to build up their reputations. For instance customers will try a new product on a small scale initially, to test for quality, and repeat purchases on a larger scale only if quality is satisfactory. Fly-by-night operators will be identified and thereafter shunned, while nonmyopic sellers will have to endure the initial testing period with low sales volumes in order to build up a reputation. The fly-by-night operators exert an externality on the nonmyopic sellers. Typically this externality is a negative one: the greater the fraction of fly-by-night operators in the population of sellers, the more severe is the sales restriction in early periods that genuine sellers will have to endure. In other words, it increases the start-up costs of getting a business successfully off the ground.

4.3 Reputation in the Indian Software Export Industry

A vivid illustration of this phenomenon is documented by Banerjee and Duflo (1999) in the context of the customized software industry in India. The product is typically complex and buyer-specific, so the quality of a delivered product is difficult to document and write into a contract that can be externally enforced. The importance of reputational considerations in getting off the ground is highlighted by the fact that it typically requires little start-up fixed capital. To start a business an entrepreneur needs merely to lease limited office space, purchase a few PCs and hire people to write software. The industry is highly competitive, being essentially driven by export orders. Monitoring development effort on the product is rendered difficult by the great distances between customer and producers. Third-party contract enforcement is further limited by the weaknesses of the Indian courts in handling business disputes speedily, and in the lack of international forums of dispute settlement. So the prospect of repeat business forms the main incentive to do a high-quality job.

On the basis of interviews of around 125 software companies in three major software development centers (Bangalore, Hyderabad and Pune) which have witnessed fast growth

rates in the recent past (exceeding 50% annually), Banerjee and Duflo find that export contracts tend to be frequently renegotiated as buyers change their requirements, and projects are subject to significant overruns. Contracts tend to be either fixed price contracts (where all the risk is on the supplier), or time-and-material contracts (where all the risk is on the buyer). Their principal empirical finding is that younger supplier firms bear more of the risk: the younger the firm the greater the likelihood of a fixed price contract, and the larger is the share of overruns borne by the supplier. Young firms (created in 1994 or later) are 26% more likely than older firms (created in 1993 or before) to have fixed-price contracts, and bear 19% more of overruns on average. They argue that these results are unlikely to be explained by efficient risk-sharing considerations, since younger firms are more likely to be wealth constrained and more risk-averse. Neither can it be explained by considerations of learning-by-doing, wherein more experienced firms encounter lower costs or deliver superior quality products, since this should affect the average price, rather than the extent of price variability that is borne by the supplier. Another evidence against the learning hypothesis is that project overruns tend to increase with the age of the firm. This is consistent with the notion that new firms are initially given ‘testing’ contracts that are less ambitious, to screen for their quality. Indeed, younger firms tend to be given contracts that were less complex and smaller in size. Firms that entered the business before 1988 had a mean project size of 190 man-months, whereas those entering after 1992 had a mean size of at most 60 man-months. The results document the empirical significance of the steep costs of building up of reputational capital for new firms.

4.4 Dual Labor Markets and Labor Productivity

One puzzle encountered by studies comparing firm productivity between developed and developing countries is in explaining large variations in labor productivity. Clark (1987) compared the efficiency of cotton textile mills in different countries of the world around 1910, that were using similar machinery. Wage rates were the highest in the USA (\$6.5-8.8 per week), followed by UK (\$5), Europe (\$2-3.8), and Japan, India and China (\$0.5-0.8). Labor costs account for over 60% of total costs, so (given data on differences in other factor

prices) they should have been decisive in determining competitiveness. In particular, the European cotton industry should have outcompeted the British industry. As it turned out, only the developing country industries in Japan-India-China (JIC) could compete effectively with the British. Clark sets out to understand this may have been the case.

The answer could not be found in differences in capacity utilization, capital-labor substitution, or in technology gaps. Clark concludes that the only possible candidate explanation lies in differences in labor efficiency: UK workers hadled an average of 2.04 looms a week, as against 1.1-1.3 in Europe, and 0.5 in JCI. These differences could not be explained by differences in education, experience, nutrition or cultural backgrounds. The only remaining possibility lie in factors specific to the location of the industry, such as climate or local labor norms, but Clark is unable to probe these factors any further.

Pack (1987) compared similar cotton mills in Kenya, Philippines and the UK, and found substantial variations in labor productivity: they were higher by 450% in the UK. These productivity variations substantially offset the wage advantage of the poorer countries: e.g., they were higher in UK by almost 9 times. But they did not offset them entirely, so the Kenyan mills were more cost-effective than the UK mills. But their competitive advantage would have been many orders higher had labor productivity been so much lower in Kenya. These differences could not be accounted for by differences in worker education, training, or differences in technology. Pack's study suggest the role of X-inefficiency and managerial slack, besides inadequate effort to screen job applicants, and the inadequate use of work incentives.

These findings are difficult to explain using conventional economic theory: if incentives increase productivity and therefore profits, why would cotton mill owners be reluctant to employ them for their managers and workers? After all, conventional moral hazard models where the only contracting constraint is the unobservability of effort, predict that incentive provisions are a property of Pareto-efficient contracts, so they cannot be viewed as an instrument of exploitation of employees by employers. If anything, the existence of a labor surplus in overpopulated developing countries suggests that employees have lower outside

options than their counterparts in developed countries, so private employers should be more able to impose incentive provisions that increase worker productivity.

Esfahani and Mookherjee (1995) provide one explanation for this phenomenon, extending the standard reputational model to a context of dual labor markets. In essence this explanation runs in terms of the informational rents necessary to provide reputational incentives to workers by employers. There is a simple moral hazard problem, wherein a hired employee can either select to work at a lower intensity, or at a higher intensity. The performance of the worker can be monitored *ex post* by the employer, but is noncontractible (as in the well known model of Shapiro and Stiglitz (1984)). So the moral hazard problem cannot be overcome by a static performance-contingent contract. Workers can be motivated only by the threat of terminating the employment contract in the future, following observation of poor performance in the current period.

Employers can choose whether or not to provide high-powered contracts that induce high effort from their employees. Consequently the labor market divides endogenously into two sectors. One sector offers short-term fixed wage contracts, which induce low effort incentives. The other offers implicit long term contracts, which pays a premium above the short-term wage to provide an incentive for workers to choose high effort. The short-term contract always clears, while the long term sector is characterized by rationing, since all workers prefer to locate in the long term contracting sector. Workers in the short-term sector constantly search for a job in the long term sector, where vacancies are created through an exogenous turnover process.

Labor productivity is clearly higher in the long term sector; with unmeasured labor quality this sector will end up with higher measured labor productivity. The composition of the industry between the two sectors therefore is a key determinant of labor productivity at the industry level: the larger the proportion of the market accounted for by the long term sector (denoted by β), the higher labor productivity will be. Unlike the models of Shapiro-Stiglitz (1984), Bulow and Summers (1986) and Eswaran and Kotwal (1985), the key question concerns the endogenous determination of β .

Firms must be indifferent between short and long-term contracts. So wage differences between the short-term (w_s) and long-term sector (w_l) must exactly offset the productivity difference ϕ : $w_s = \phi w_l$. This implies that the *absolute* difference between the two wages $w_s - w_l$, or the wage premium from the worker point of view, is increasing in the average wage. In a labor surplus economy, the average wage will be small, and hence so must be the wage premium be. Everything else remaining the same, this lowers the incentive role of termination threats in the long term sector. To compensate for this, it must be the case that it is more difficult for a discharged worker to find another long-term job quickly. In other words, β must be lower: jobs in the long-term sector must be few and far between. So labor surplus economies are inherently prone to lower labor productivity, which partially offset their natural advantage from their labor abundance, exactly as the studies of Clark and Pack documented.

Similar results are obtained for employers' incentives to provide on-the-job training to their workers by Majumdar (1998) in a related model. Given that workers are capital-constrained, employers must invest in their training. Then after being trained workers can hold-up the employer *ex post* by threatening to work for other employers, if the training involves some element of general rather than firm-specific skills. Credible outside offers are determined endogenously as the result of bargaining between the worker and outside firms. In the model the extent to which holdup threats by the worker are effective depends on the the *ex post* to the firm of replacing a trained worker by a suitable number of untrained workers from a competitive spot market. In a labor surplus economy, the the wage cost of untrained workers is low, which tends to reduce the extent of holdup. A countervailing effect on *ex ante* training incentives is that the benefit of training workers is lower, since a trained worker can be replaced by a number of untrained workers. The latter effect turns out to dominate, so that a labor surplus economy generates lower incentives to firms to train their workers.

Both models predict that the entry of new firms in the industry will tighten the labor market, tending to raise the general wage level, and thus increase the size of the long term sector. In this sense increased competition in the industry will boost labor productivity. An

expansion into foreign markets will have a similar effect, so measures to raise exports will automatically improve labor quality. Measured competitiveness prior to entering foreign markets will be a poor guide to predicting their eventual success. However, protecting the home market to foreign competition will also have the same effect, so these considerations suggest the efficacy of policies that limit competition. In general, at the industry level the externality effect operates similar to a Marshallian scale economy.

4.5 Subcontracting Networks: Matching Demand Fluctuations with Capacity

Finally, we turn to subcontracting networks that have been observed in many instances of successful industrial clusters of small firms, such as the footwear industry in Brazil, Italy and Mexico, or the garment industry in South India or Bangladesh (described in the 1995 symposium issue of *World Development*). Humphry (1995) describes these industrial districts in terms of four key factors:

a cluster of small and medium enterprises spatially concentrated and specialized in specific industries

forward and backward linkages between these enterprises, based on both market and non-market exchanges of goods, information and people

a common cultural and social background linking these agents, and the use of implicit agreements concerning exchanges that are upheld via social norms

public and private local institutions (e.g., to promote technological exchanges and commercial promotion) acting to support the cluster

The basic principle is that

“...clusters of predominantly small firms can gain economies of scale and scope and increased flexibility through specialization and interfirm competition. If

they cluster they can be as competitive or more competitive than larger firms.”

(Humphrey (1995, p. 1))

Kranton and Minehart (1998) develop a theory of such networks as an alternative to markets and vertically integrated enterprises. Consider a manufacturing process which produces a finished product through two successive stages of production. In their framework, *markets* refer to the purchase of nonspecialized intermediate good by the downstream manufacturer on a competitive spot market. *Vertical integration* involves the downstream producer integrating backwards to produce capacity for specialized intermediate goods. The use of a specialized intermediate improves the quality of the product, as well as the reliability of input supply. However, this requires a substantial capital investment by the downstream manufacturer. Also, when demand for the finished product is variable, the capacity for producing the intermediate good is frequently underutilized.

Finally, *networks* are represented by links between multiple downstream and upstream enterprises. The upstream enterprises produce specialized intermediate goods for downstream enterprises that invest in creating such a link. Relative to the market, the downstream producers achieve a quality improvement owing to the use of specialized inputs. It also economizes on the need for downstream enterprises to invest in the upstream process, e.g., compared with vertical integration. The pattern of exchanges between the downstream and upstream enterprises reacts flexibly to *ex post* demand shocks faced by different downstream firms, either through spot auctions or other allocation mechanisms. The cost, however, relative to vertical integration, is that downstream buyers face the risk of being rationed *ex post* in the allocation of their specialized input. An added complication, not explicitly analysed in their model, concerns problems of quality assurance in the specialized inputs supplied by the upstream firms, which are frequently overcome through long-term dealings and shared social and cultural ties.

Kranton and Minehart compare the level of welfare resulting from the three distinct organizational forms, and compare these with noncooperative equilibrium outcomes. Networks dominate markets when quality improvements resulting from specialized inputs are

significant. They dominate vertical integration when demand for finished goods is more volatile, and the costs of investing in upstream capacity are large. This explains the importance of such networks in garment and footwear industries where quality and demand volatility are especially important. Moreover, they tend to be frequently observed in developing countries where (i) greater capital market imperfections make it more difficult for downstream firms to raise the capital necessary to integrate backwards; and (ii) community networks and informal cooperation play a more significant role in economic activity.

Banerjee and Munshi (1999) empirically study the garment export network in the southern Indian town of Tirupur, which has achieved striking success since the 1980s. Tirupur is a small town which accounts for over 50% of India's garment exports, and also supplies a significant fraction of the Indian market. Exports have grown at annual rates exceeding 50% since 1985, while the town's population doubled between 1971 and 1991. Similar to the industrial clusters in the footwear industry in Italy, Brazil and Mexico, much of the Tirupur industry involves significant subcontracting between different stages of production. This is particularly the case for one community, the Gounders, which dominates the industry. Other communities which have migrated to Tirupur from different parts of Northern and Western India in contrast tend to invest more in vertically integrated enterprises.

Banerjee and Munshi consequently compare the performance of enterprises in these different communities, which illustrate the comparative advantages of networks *vis-a-vis* vertical integration, besides the role of communities in the industrialization process. The main differences they identify are: (i) the Gounders have access to an intracommunity subcontracting network, which tends to reduce the significance of hold-up problems owing to shared kinship and cultural ties, whereas the outside communities do not have access to this network; and (ii) the two communities operate in two different capital markets. The Gounders accumulated substantial capital from cash crop cultivation in the earlier part of the 20th century, and have few profitable outlets for investing their surpluses outside Tirupur. In contrast the outsiders have access to investment opportunities outside Tirupur, and are more reluctant to sink their investments in immobile assets in Tirupur.

The consequence is that the Gounders invest greater amounts of capital in their enterprises. Banerjee and Munshi argue that this played a significant role for the Gounder community in achieving higher export growth rates, based on an econometric analysis of exports, investments and finance of 600-odd individual enterprises.

5 Conclusion

I have argued that many distinctive characteristics of industrial organization in developing countries can be understood as a consequence of moral hazard problems and impediments to the resolution of these problems via contractual means. Three kinds of impediments were discussed: financial constraints, inability to write complete long term contracts, and problems of contract enforcement. They respectively give rise to three alternative ‘solutions’, involving contractual payment of ‘informational rents’, reallocating property rights, and relying greatly on reputational considerations. All of them raise the costs of dealing with the moral hazard problem, creating an ‘institutional’ source of lack of competitiveness of firms in developing countries. They are embodied in greater proneness to imperfections in various factor markets, chiefly capital, material inputs, and labor. Combined with poorer infrastructural services, these imperfections raise the costs of producing high quality, reliable products that can effectively compete with firms in more developed countries.

The chief determinants of ‘institutional’ comparative advantage include the nature of (i) wealth and income inequality, and nature of financial institutions, which affect the proneness of entrepreneurs to capital shortages, both physical and human; (ii) community networks, which determine the extent to which these capital shortages can be overcome via institutions of exchange and reciprocity between different stages of vertically disintegrated production; (iii) infrastructure services, especially their quality and reliability; and (iv) agricultural prosperity and nature of agrarian relations, which determine the volume of investible surpluses in industry, as well as contracting relationships between agroprocessing industries and their crop suppliers.

These contracting imperfections also explain the importance of distinct organizational forms that are observed in the developing world, such as cooperatives, joint ventures and community-based clusters of subcontracting relationships. Successful industrialization often requires a clustering of numerous favorable institutional attributes, such as communities with strong internal networks, large investible surpluses, aggressive entrepreneurial attitudes, and access to reliable public infrastructure. Otherwise it is based on large scale conglomerates controlled by a few wealthy families that are horizontally and vertically integrated, with close links to the government to assure access to reliable infrastructure and finance. This necessarily imparts a certain unevenness in the patterns of industrialization. Entrepreneurs that do not belong either to wealthy families or to business/industrial communities will find it hard to start and grow a business. In turn this implies that the benefits of industrial progress will be less widely dispersed compared with developed countries, excluding a large majority of the population.

Since many of the key imperfections involve factor markets, which are less easy to ‘liberalize’ and ‘deregulate’ than are product markets, the problems of finance and of reliable input supply persist, despite substantial market-oriented reforms in recent years in many developing countries. As Khanna and Palepu (1998) argue, this accounts for the persistence of large business conglomerates in Chile and India in the 1980s and the 1990s. The contracting impediments are less easy to get rid of by purposive policy measures of deregulation, and are more thus more likely to come to the fore. The kinds of policies that can make a dent on the factor market imperfections — improvements in human development, financial institutions and social capital — are notoriously difficult to implement and take a long time to take effect. Until then developing countries are unlikely to realize their full potential to benefit from forces of globalization.

References

Akerberg Daniel and Maristella Botticini (1999), “Endogenous Matching and the Empirical Determinants of Contract Form,” Discussion Paper n0. 92, Institute for Economic Development, Boston University.

Aghion Phillipe and Patrick Bolton (1997), “A Trickle Down Theory of Growth and Development,” , *Review of Economic Studies*, 64, 151-172.

Akerlof George (1970), “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics*, August, 488-500.

Amin Shahid (1984), *Sugarcane and Sugar in Gorakhpur*, New Delhi: Oxford University Press.

Banerjee Abhijit and Andrew Newman (1993), “Occupational Choice and the Process of Development,” *Journal of Political Economy*, 101(2), 274-298.

————— (1994), “Poverty, Incentives and Development,” *American Economic Review, Papers and Proceedings*, 84(2), 211-215.

————— and Esther Duflo (1999), “Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry,” mimeo, Department of Economics, MIT.

————— and Kaivan Munshi (1999), “Market Imperfections, Communities and the organization of Production: An Empirical Analysis of Tirupur’s Garment Export Network,” mimeo, Department of Economics, MIT.

—————, Dilip Mookherjee, Kaivan Munshi and Debraj Ray (1997), “Inequality, Control Rights and rent-Seeking: Sugar Cooperatives in Maharashtra,” Discussion Paper No. 80, Institute for Economic Development, Boston University.

Bardhan, Pranab (1989), (Editor), *The Economics of Agrarian Institutions*, Clarendon Press; Oxford.

Bester Helmut (1985), "Screening vs. Rationing in Credit Markets with Imperfect Information," *American Economic Review*, 75(4), 850-855/

Bowles Samuel and Herbert Gintis (1994), "Efficient Redistribution: New Rules for Markets, States and Communities," Working Paper 1994-7, Department of Economics, University of Massachusetts, Amherst.

————— (1995), "The Distribution of Wealth and the Assignment of Control Rights," mimeo, Department of Economics, University of Massachusetts, Amherst.

Braverman Avishay and Joseph Stiglitz (1982), "Sharecropping and Interlinking of Agrarian Markets," *American Economic Review*, 72, 289-312.

Bulow Jeremy and Lawrence Summers (1986), "A Theory of Dual Labor Markets with Applications to Industrial Policy, Discrimination and Keynesian Unemployment," *Journal of Labor Economics*, 1(3), 376-414.

Cawthorne Pamela (1995), "Of Networks and Markets: The Rise and Rise of a South Indian Town, the Example of Tiruppur's Cotton Knitwear Industry," *World Development*, 23(1), 43-56.

Chithelen Ignatius (1985), "Origins of Co-operative Sugar Industry in Maharashtra," *Economic and Political Weekly*, Vol. XX, No.14.

Clark Gregory (1987), "Why Isn't the Whole World Developed? Lessons From the Cotton Mills," *Journal of Economic History*, 47(1), 141-173.

Coate Stephen and Martin Ravallion (1993), "Reciprocity without Commitment: Char-

acterization and Performance of Informal Insurance Arrangements,” *Journal of Development Economics*, 1-24.

Dutta Bhaskar, Ray Debraj and Kunal Sengupta (1989), “Contracts with Eviction in Infinitely Repeated Principal-Agent Relationships,” in *The Economic Theory of Agrarian Institutions*, P. Bardhan (ed.), Clarendon Press, Oxford.

Eaton Jonathan and Mark Gersovitz (1981), “Debt with Potential Repudiation: Theoretical and Empirical Analysis,” *Review of Economic Studies*, 289-309.

Esfahani Hadi (1991), “Reputation and Uncertainty: Toward an Explanation of Quality Problems in Competitive LDC Markets,” *Journal of Development Economics*, 35, 1-32.

——— and Dilip Mookherjee (1995), “Productivity, Contracting Modes and Development,” *Journal of Development Economics*, 46, 203-231.

Eswaran Mukesh and Ashok Kotwal (1985), “A Theory of Two-Tier Labor Markets in Agrarian Economies,” *American Economic Review*, 75, 162-177.

Foster James and Henry Wan (1984), “Involuntary Unemployment as a Principal-Agent Equilibrium,” *American Economic Review*, 74(4), 476-484.

Galor Oded and J. Zeira (1993), “Income Distribution and Macroeconomics,” *Review of Economic Studies* **60**(1), 35–52.

Ghatak Maitreesh, Massimo Morelli and Tomas Sjostrom (1998), “Dynamic Incentives, Occupational Mobility and the American Dream,” mimeo, Department of Economics, Iowa State University.

Ghosh Parikshit and Debraj Ray (1996), “Cooperation in Community Interaction without Information Flows,” *Review of Economic Studies*, 491-519.

————— (1999), “Information and Enforcement in Informal Credit Markets,” Discussion Paper No. 93, Institute for Economic Development, Boston University.

Grossman Sanford and Oliver Hart (1983), ‘An Analysis of the Principal Agent Problem,” *Econometrica*, 51, 7-45.

————— (1986), “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration,” *Journal of Political Economy*, 94(4), 691-719.

Hart Oliver and John Moore (1990), “Property Rights and the Nature of the Firm,” *Journal of Political Economy*, 98, 1119-1158.

————— (1994), “The Governance of Exchanges: Members Cooperatives vs. Outside Ownership,” Harvard University and London School of Economics, mimeo.

Hoff, Karla (1994), “The Second Theorem of the Second-Best,” *Journal of Public Economics*, 45, 223-242.

————— and Andrew Lyon (1994), “Non-Leaky Buckets: Optimal Redistributive Taxation and Agency Costs,” *Journal of Public Economics*,, 45.

————— and Joseph Stiglitz (1998), “Moneylenders and Bankers: Price Increasing Subsidies in a Monoplistically Competitive Market,” *Journal of Development Economics*, 55, 485-518.

Holmstrom Bengt (1979), ‘Moral Hazard and Observability,” *Bell Journal of Economics*, 10, 74-91.

————— and Jean Tirole (1994), “Financial Intermediation, Loanable Funds and the Real Sector,” Working Sector, IDEI, University of Toulouse.

Humphry John (1995), "Introduction," *World Development*, 23(1), 1-7.

Kahn Charles and Dilip Mookherjee (1995), "Market Failure with Moral Hazard and Side Trading," *Journal of Public Economics*, 58, 159-184.

————— (1998), "Competition and Incentives with Non-Exclusive Contracts," *Rand Journal of Economics*, 29(3), 443-465.

Khanna Tarun and Krishna Palepu (1998), "Policy Shocks, Market Intermediaries and Corporate Strategy: The Evolution of Business Groups in Chile and India," mimeo, Harvard Business School.

Klein Benjamin, R. Crawford and Armen Alchian (1978), "Vertical Integration, Appropriate Rents and the Competitive Contracting Process," *Journal of Law and Economics*, 21, 297-326.

Klein Benjamin and Keith Leffler (1981), "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89, 615-641.

Kranton Rachel (1996), "The Formation of Cooperative Relationships," *Journal of Law, Economics and Organization*, 214-233.

————— and Deborah Minehart (1998), "Vertical Integration, Networks and Markets," mimeo, Department of Economics, Boston University.

Laffont Jean-Jacques and M. Matoussi (1995), "Moral Hazard, Financial Constraints and Sharecropping in El Ouja," *Review of Economic Studies*, 62(3), 381-400.

Legros Patrick and Andrew Newman (1996), "Wealth Effects, Distribution and the Theory of Organization," *Journal of Economic Theory*.

Lin C-C. and Ivan Png (1998), "Kinship, Control and Incentives," mimeo, Department of Economics, National Chengchi University, Taiwan.

Ljungqvist Lars (1993), "Economic Underdevelopment: The Case of a Missing Market for Human Capital," *Journal of Development Economics*, 40, 219-239.

Loury Glenn (1981), "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49, 843-867.

Lucas, Robert (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1), 3-42.

Majumdar, Sumon (1998), "Underinvestment in Worker Training and Underdevelopment," mimeo, Department of Economics, Boston University.

Mirrlees James (1975), "The Theory of Moral Hazard and Unobservable Behavior - I", mimeo, Nuffield College, Oxford.

Mookherjee Dilip (1997a), "Informational Rents and Property Rights in Land," in *Property Rights, Incentives and Welfare*, J. Roemer (ed.), Macmillan Press, forthcoming.

————— (1997b), "Wealth Effects, Incentives and Productivity," *Review of Development Economics*, 1(1), 116-133.

————— and Debraj Ray (1998), "Persistent Inequality and Endogenous Human Capital Investment Thresholds," mimeo, Department of Economics, Boston University.

————— (1999), "Contract Structure and Wealth Accumulation," mimeo, Department of Economics, Boston University.

Newman Andrew (1995), "Risk Bearing and the Knightian Theory of Entrepreneurship," mimeo, Department of Economics Columbia University.

Pack Howard(1987), *Productivity, Technology and Industrial Development*, New York: Oxford University Press.

Piketty Tomas (1997) "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing," *Review of Economic Studies*, 64, 173-189.

Riley John (1987), "Credit Rationing: A Further Remark," *American Economic Review*, 77(1), 224-227.

Sawhney Dhruv (1997), 'The Indian Sugar Industry,' mimeo, International Society of Sugar Cane Technologists, New Delhi.

Schultz Theodore (1968), "Institutions and the Rising Economic Value of Man," *American Journal of Agricultural Economics*, 50, 1113-1122.

Shapiro Carl (1983), "Premiums for High Quality Products as Returns on Reputation," *Quarterly Journal of Economics*, 98, 659-680.

——— and Joseph Stiglitz (1984), "Worker Unemployment as a Worker Discipline Device," *American Economic Review*, 433-444.

Shetty Sudhir (1988), "Limited Liability, Wealth Differences and Tenancy Contracts in Agrarian Economies," *Journal of Development Economics*, 29(1), 1-22.

Stiglitz Joseph and Andrew Weiss (1981), "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71, 393-419.

Thiele Henrik and Achim Wambach (1998), "Agency Costs and Wealth Effects in the Principal Agent Model," mimeo, Department of Economics, University of Munich.

Tirole Jean (1996), "A Theory of Collective Reputations," *Review of Economic Studies* 63.1, 1-22.

Williamson Oliver (1975), *Markets and Hierarchies*, Free Press, New York.

————— (1985), *The Economic Institutions of Capitalism*, Free Press, New York.