SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

The Case of the Errant Executive : Management, Control and Firm Size in Corporate Cheating

Brishti Guha

September 2005

Paper No. 16-2005

The Case of the Errant Executive : Management, Control and Firm Size in Corporate Cheating

Brishti Guha[1]

Singapore Management University

Abstract

Firm insiders – a manager and a board – face moral hazard in relation to their outside shareholders in a repeated game with asymmetric information and stochastic market outcomes. The manager determines whether or not outsiders are cheated; the board, whose objectives differ from those of outside shareholders, attempts to control the manager through compensation contracts and dismissal threats Since compensation determines the manager's incentive to cheat, firms competing for outside capital publicly announce their managerial contracts. However, secret renegotiation between firm and manager is still possible: so outsiders guard against being cheated by limiting their total stake in any firm. This imposes a *credibility constraint* on firm size, providing a rationale for the shape of long-run cost curves. Given this limit on outside funds, the minimum size requirement for enterprises to become operational and the ability to pay managers enough to ensure honesty both set a floor to the personal wealth required to enter entrepreneurship. Thus, we endogenize entry into industry, establish a unique equilibrium for any distribution of wealth, and characterize different equilibria. We also explain features of poor countries like dominance of family firms, moral hazard induced vicious circles that retard industrialization and the stimulus that inequality *may* provide to industrial development.

Keywords: Moral hazard, firm size, managerial compensation, repeated games.

JEL Classification: D82, L21, M52, O11.

---

[1] School of Economics and Social Sciences, Singapore Management University, 90 Stamford Road, Singapore 178903. E-mail : bguha@smu.edu.sg . I began work on this paper while a PhD student at Princeton University.

1.Introduction

In this paper we analyze a setup where firm insiders – a board and a manager – face moral hazard with regard to their outside shareholders. While the manager decides whether or not outside shareholders are cheated, the board – whose interests do not coincide with the manager's or the outside shareholders' – may seek to influence him through instruments like the managerial compensation contract and threats of dismissal. We argue that public observability of managerial contracts does not in itself eliminate moral hazard, given the possibility of secret renegotiation ; and that outsiders can work out a "credibility ceiling" on the ratio of outsider to insider financing consistent with managerial honesty. This pins down firm size in equilibrium, thus providing one possible answer to the long-standing controversy about long-run limits to firm size, and the shape of long-run cost curves.

Given our credibility constraint, two forces impose a floor on personal wealth required to enter entrepreneurship. The first is a minimum size requirement on enterprise. As credibility limits the amount of outsider finance forthcoming, the personal wealth of potential entrepreneurs has to exceed a certain floor. The second is that entrepreneurs must pay managers enough to ensure honesty – offering them "efficiency wages" of the Shapiro-Stiglitz variety (perhaps one should call them "honesty wages" in this context). This enables us to endogenize entry into entrepreneurship and characterize possible equilibria (there is a unique equilibrium for a given distribution of initial wealth). Apart from implications for firm size, our model also explains features of poor countries such as the prevalence of family firms, moral hazard induced "vicious circles" retarding industrialization, and the possible advantages of inequality in the industrial development of a small open economy. The link with the realm of development economics lies in the fact that our equilibria are influenced by the level and the distribution of wealth.

Market outcomes in our model are stochastic, and the true performance level of a firm is only observable by insiders. Insiders can therefore exploit their asymmetric information about firm performance to pretend to outsiders that the firm has been unlucky, appropriating the excess returns due to outsiders in a good state. We explicitly introduce an incomplete contracting feature in the agreement between insiders and outside investors, due to the inability of the latter to observe firm performance.

Investor payoffs are common knowledge, but this does not tell the public whether the insiders have cheated. This is because outcomes being stochastic, poor investor returns may reflect either bad luck or cheating by insiders. This is similar in some sense to the Green-Porter (1984) models of games with imperfect public information[2].

The manager has the ability to cheat outside investors in the manner described. A public signal detects cheating with a probability of q – but only in the next period. Past cheating by a firm, once exposed, is collectively remembered for ever.

Our work is connected to many strands in the literature. Agency conflicts between managers and shareholders have of course been the subject of a vast literature in corporate finance (eg. Aghion and Bolton [1992], Jensen [1986], and Zwiebel [1996]). All these papers consider managers whose objectives differ from those of shareholders, or of investors in general. The difference in objectives may be reflected in the manager's choice of projects inimical to investor interests, or in his diverting free cash flow into his own pockets. The latter option is closer to the situation we study in this paper. Our focus is on a three-way relationship between boards, managers and outside investors, rather than only on manager-shareholder conflicts.

Our paper is also related to the literature on incomplete contracting and non-verifiability. This includes work on "relational contracting" (eg Greif [1991,1996]) and on honesty versus cheating in repeated game setups (eg Dixit [2003]). Dixit considers a prisoner's dilemma between randomly matched players. In his paper, observability of payoffs is a sure guide to whether or not cheating has occurred – but this is not the case in our model. In ours, in spite of perfect observability of investor payoffs not just by the cheated investor but by the general public, extraneous uncertainty regarding market outcomes means that even the cheated party cannot tell for sure (without an imperfect lagged public signal) if it has really been cheated. Unlike in the Dixit paper, in ours

[2] Green and Porter (1984) discuss how low sales may reflect either low demand or rival pricing strategies by other firms.

players are not randomly matched – investors have the choice of continuing with the same firm, and boards have the option to retain the same manager in future.

Literature on controversies surrounding the long-run cost curve and the determinants of long run firm size is also relevant. This dates back to Kaldor's [1934] argument that a determinate limit on firm size within a particular industry requires a factor that is in limited supply to the individual firm, even in the long run, but has a definite supply price for the industry as a whole. Kaldor argued that long run firm size is in fact indeterminate, contesting the view that "co-ordinating ability" was in fact such a factor and that costs of co-ordination limited firm size in long-run equilibrium. Whether or not we accept Kaldor's views (elaborated on later in this paper), our paper provides a different mechanism from the standard textbook view to pin down firm size. Credibility concerns caused by the possibility of secret renegotiation between the board and the manager affect the behavior of outside investors, so that the supply of outside capital to each firm is limited, even in the long run, while it has a definite supply price – derived in our paper – for the industry as a whole. Our explanation of the limits on firm size in the long run relies not on technology – broadly defined as properties of production functions - but on the presence of moral hazard.

Our paper also relates to the development literature on imperfect markets for capital or credit. For example, Banerjee and Newman (1993) trace occupational choice to the initial wealth distribution. Credit is rationed to entrepreneurs on the basis of collaterizable wealth while each enterprise has a minimum investment requirement. So only those whose wealth exceeds a certain floor can become entrepreneurs. In our paper we combine indivisibilities in firm size with the moral hazard that entrepreneurs (insiders in our model) face with regard to their outside shareholders. We have already mentioned that the idea of paying a manager substantially more than his outside option, which happens in our model, is akin to the Shapiro-Stiglitz (1984) "efficiency wage" concept in the context of imperfect monitoring.

Our paper is also related to empirical work on corporate cheating and governance. Joh (2003) and Lemmon and Lins (2003) find, using evidence from East Asia, that malfeasance tends to be higher in firms with a divergence between ownership and control – that is, a high ratio of outsider to insider financing. This supports our theoretical prediction that a high ratio of outsider to insider financing intensifies moral hazard. Peng and Roell (2004) show that the number of investor lawsuits brought against firms (i.e. the incidence of cheating) is negatively and significantly affected by the ratio of the "fixed salary" to the "bonus" component of executives' compensation. This fits in with our result that cheating is better deterred by paying the executive a fixed salary rather than a combination of a salary and a share in profits and that a high enough salary component can ensure honesty. Other empirical work in support of this includes Graef Crystal's (1991) finding that "long-term firm performance" as measured by ten-year shareholder returns (accounting for dividends as well as stock-price appreciation) is negatively and significantly related to the "long-term incentive payments" to the firms' executives (these were in many forms and could involve giving the executive stocks, options, etc). If poor long-term performance reflects executive cheating, then a larger component of share-based pay seems to be positively correlated with cheating.

This brings up the issue of profit-based executive compensation. Many have pointed out the incentive effects of such compensation, arguing that bonuses induce greater efficiency. However in our story, output, though stochastic, does not depend on the manager's effort but on luck. So the effect of giving the manager a share in insider profits is to increase his incentive to cheat the investors by exploiting his asymmetric information about whether the firm has experienced good or bad fortune.

Anderson and Bizjak (2003) find that more independence for the compensation committee over time does not seem to have reduced executive pay. One explanation in terms of our model would be that managers need to be paid a certain amount to remove their incentives to cheat. This payment must be substantially greater than the managers' outside opportunity cost (which for simplicity has been assumed here to be zero).

In section 2, we provide a detailed discussion of the assumptions underlying our model. In section 3, we characterize the managerial contract given public observability of managerial compensation while in section 4, we derive the effect of possible renegotiation. In section 5, we discuss the "credibility ceiling" and show that moral hazard imposes two constraints on the personal wealth of would-be entrepreneurs and on firm size; we also characterize possible equilibria for different distributions of initial wealth. In section 6 we discuss some implications of our model, while section 7 concludes.

2. Assumptions

Individuals live forever in a closed population with fixed and inelastic wealth. There is no saving[3], depreciation or borrowing, though lending is possible at an outside opportunity cost R. Agents are all risk-neutral.

Each enterprise must reach a minimum size $\underline{I}$ before operation is possible. Thus for entrepreneurs whose funds F are less than this, going public is a necessity to set up business. F may represent the collective funds of a group of insiders who combine to set up a firm. Collective action problems limit the extent to which such pooling of funds is possible so that the minimum size requirement has considerable bite.

Each enterprise lasts for one period after which outside investors have to decide whether to refinance the enterprise, or to costlessly transfer their capital to some other enterprise, or to invest it in the outside option.

Enterprises earn a rate of profit G with probability p (the probability of "good luck"), and B otherwise, with $G > B$. The expected rate of profit is $H = pG + (1 - p)B$ and exceeds R, justifying the existence of the industry. Here G and B are exogenous - we are operating in a small open economy facing fixed world prices and exogenous shocks to output.

In the 'understanding' between insiders and outsiders, outsiders are to receive an expected dividend of D (determined endogenously by the wealth distribution) on their capital S, the actual amount received in any period being proportional to G/H or B/H depending on whether good or bad luck has been realized. Insiders as a whole are to receive an expected dividend of D on their capital F plus an extra amount such that the insiders' and outsiders' payoffs sum to total firm profits. Insiders on average must get at

[3] This assumption is discussed later in this section.

least as much as outsiders do, otherwise they prefer to be investors in other people's firms rather than to set up firms of their own. To simplify the analysis, we assume that this is true not only ex ante, but ex post as well. The firm's assets are redeployable so that if the enterprise is dissolved, erstwhile insiders may still invest their assets in other firms or in the outside option.

The understanding between insiders and outsiders is not enforceable because the state of firm performance (whether the firm has experienced good or bad luck) is observable only by insiders - it is not observable or legally verifiable by outsiders. Insiders can cheat by paying the outsiders their "bad luck" dues even when good luck has occurred – appropriating the excess. F,S,G,B and p are all taken as exogenous by the individual outsider. Moreover, outsiders can observe the aggregate ratio of outsider to insider capital, which we denote by s.

A publicly observable signal detects cheating with an accuracy (probability) q but only after it has occurred. The information it conveys is available to all investors at the beginning of the next period.

We assume that $DB/H < R$, so that investors prefer not to enter the industry if they expect to receive only their bad luck dues. A sufficient condition for this to hold for all non-negative D and positive R is that B should be non-positive. A weaker sufficient condition is that B should be smaller than R. This is sufficient because as we will argue below, D can never credibly exceed H as this would be known to violate the insider's participation constraint, making potential insiders more eager to become outside investors in other firms than to set up their own firms.

We also introduce an agency structure within the firm. We assume that the firm is characterized by an internal division of labor - all executive decisions are taken by an executive[4] with his personal objective function while the "firm" is broadly defined as the authority that hires and fires him. The firm can use two instruments to control the executive – the compensation contract (assumed to be perfectly verifiable and enforceable) and the threat of dismissal (where this threat is credible). The executives unconditionally

[4]We use the terms "executive" and "manager" interchangeably.

maximize their expected payoff. We assume further that the supply of executives is infinitely elastic at price zero – they do not in effect have an outside option

Traditionally the agency problem has been viewed as a conflict between the interests of the shareholders and the manager. A vast body of literature in corporate finance (Aghion and Bolton(1992), Innes (1990), Jensen(1986) etc) deals with a manager who either because of his private objectives or due to the nature of his compensation contract, may take decisions which are sub-optimal for the firm – whether such decisions involve exerting too little effort, or choosing the wrong kinds of projects. In our treatment of the theme, we emphasize that the "investors" or "outside shareholders" are distinct from the "firm insiders" while at the same time, we also stress a possible conflict of interests *within* the insiders - the board and the executive. The distinction we draw between insiders and outsiders is that only insiders, who directly run the firm, know the true state of firm performance. We also assume that the outside shareholders do not have an active say in decisions such as hiring and firing managers, which is typically something the board of directors would do. However the outsiders can observe such actions taken by the board and may draw their own inferences. Thus in our setup shares may be too widely dispersed for individual outside shareholders to be able to effectively exercise control over management.

As for the conflict of interest between the manager and the board, the manager can siphon off private gains for himself beyond those envisaged by the board while setting his compensation package. Thus, a manager, while cheating investors, can also appropriate some "private gain" ε (subject to a cap[5]) which comes out of the firm. We assume that the board cannot prevent this or that the extraction of this ε is not verifiable in courts and so cannot be punished by law though the board can try to control the manager using the instruments outlined above. Where these instruments cannot serve as effective threats, the conflict becomes more marked. The introduction of private gains drives a possible wedge between the board's interests and the manager's.

[5] The logic for a cap, as explained later in the text, is that appropriation of very large gains by the manager may be easier to prove in court – verifiability beyond a limit implies that appropriation of gains beyond this cap can be ruled out by sufficiently harsh penalties.

Can the "manager" we model below be considered the CEO of the firm? The manager in our model takes the action which determines whether or not insiders as a whole cheat outside shareholders. He also has a chance to steal directly from other insiders. Whether he will in fact do this will depend on the contract offered and his personal optimization decision. Now one of our assumptions here is that the board determines the compensation of the manager. This manager could then be the CEO if we believe that the board of directors is independent of the CEO and can make autonomous decisions regarding his compensation, his appointment or his dismissal. While this is true in theory, case studies, such as Graef Crystal's, indicate that the CEO often suggests the salaries that board members should pay themselves. This introduces circularity in the sense that the board's decisions regarding the CEO are then affected by an added factor, the influence of the CEO on the salaries and perks of its members. Moreover, sometimes CEOs have served on the compensation committees of their own boards. These facts seem to suggest that an analysis which draws a distinction between the executive and the board could be more aptly applied to some other executive who does not have this kind of power over the board's decisions. In recent times, however, the compensation committees of boards have become much more independent. As Anderson and Bizjak (2003) point out, changes in SEC rules in 1992 and 1993 discouraged senior executives from serving on compensation committees and by the end of 1998 CEOs were essentially absent from compensation committees in the 110 NYSE firms in their sample. Recent evidence also points to the fact that CEOs no longer serve on compensation committees. Further, we do not rule out collusion to cheat investors between board and manager. As for hiring and firing, in theory the board has the right to hire and fire CEOs, but cases abound of the CEO appointing his cronies to the board – even if they are not insiders they may have some sort of link with the CEO. These would fall under the purview of the collusion alluded to above. Moreover, though we assume that the board can fire the executive, we predict that such firing would not occur in equilibrium. Empirically this would lead us to observe that senior executives or CEOs are only fired very rarely.

Of our assumptions, the one that needs further discussion is that of no savings. We assume zero savings so as to focus on the distinctive consequences of cheating without our results being obscured by the changing dynamics of the accumulation and distribution of wealth. Zero saving makes the distribution of wealth exogenous. It eliminates complex feedback effects such as the possibility that firms which go public may ultimately save enough out of current income to raise entrepreneurial wealth above the threshold needed to set up business without reliance on outside capital.

Somewhat extreme assumptions with regard to savings are in fact common in the literature where they serve a variety of purposes. For example, Galor and Zeira [1993] assume that no one consumes at all in the first of the two periods that he lives; Bernanke and Gertler [1989] assume that some agents consume only after retirement. Some licence is usually permitted in the literature with regard to the savings assumption.

In our model, however, the no-savings assumption can possibly be justified as follows. With the standard postulate of risk-neutrality and constant time-preference, the intertemporal utility function can be written as

$$U = \sum \delta^t c_t$$

where c_t is consumption in the t-th period. The net gain in utility from a one-period postponement of a unit of t-th period consumption is then

$$\delta^t[-1 + \delta(1 + r_t)]$$

where r_t is the return to capital in the t-th period. With risk-neutrality, savings are no longer needed to smooth consumption. They now reflect only the difference, if any, between the rates of time-preference and of return to capital. When these are independent of consumption, savings have a bang-bang character. If capital is consumable and time preference higher than the rate of return, all wealth is consumed in the first period. If the rate of return is higher, all income is saved and consumption perpetually deferred . Savings will be exactly zero if (1) capital is not consumable (again a standard assumption, see Bernanke and Gertler, 1989) and (2) time preference is higher than the rate of return.[6] In a model where the highest rate of return is H, it is sufficient for zero savings if $H < (1 - \delta)/\delta$ – a restriction not inconsistent with any of our results.

[6] No individual can dissave by trading capital for output, since, if one wishes to dissave, so will everyone else – so that the potential dissaver cannot find anyone to trade with.

One question of course remains. Where did the wealth come from if there are no savings? All wealth could be land, where output has a life-span of just one period. Alternatively, in an industrial economy, wealth could be machinery, that the country has received through foreign aid or as war reparations. We wish to focus on the problem of cheating independently of the level or distribution of wealth; and all we need is that the zero-savings assumption should be self-consistent, not that it should be realistic.

3. The Honesty Prison

Managerial contracts are common knowledge when initially announced. A contract (Φ, A^*) specifies the manager's share Φ in the insiders' profits as well as a fixed salary A^*. Φ includes a share not only in the insiders' legitimate profits, but also in their one-time cheating gains – extracted by paying outsiders their "bad luck dues" even when luck has been good. The size of such gains would then be the difference between the outsiders' dues in the good and the bad states – that is, $DS\frac{G}{H} - DS\frac{B}{H} = \frac{DSL}{H}$ where $L = G - B$. We focus on contracts that specify non-negative values of both Φ and A^* and a positive value for at least one of them**.** In addition to his contractual income, the executive can steal an amount ε from the firm. ε is capped at $\bar{\varepsilon}$ - perhaps because a larger theft would be provable in a court of law, and therefore may be ruled out by sufficiently harsh penalties. All thefts, in turn, run the risk of being detected (with a probability q) by the public signal at a later date.

The timing of moves is as follows. First, firms announce compensation contracts. Boards recruit from the pool of available executives. Then outside investors decide whether to invest or not. The information available to them at this point relates to the managerial contract, the firm's choice of managers, the payoffs distributed by the firm earlier and the public signal that gives a clue as to whether the insiders cheated in the previous period. Finally, firms realize their outcomes. Managers distribute payoffs to shareholders, having decided whether to be honest or to cheat (their outside shareholders and also perhaps their boards) in the process. The public signal indicates cheating by the executive with probability q. Firms decide whether to retain their executives or to dismiss them. Investors decide whether to reinvest in their existing firm or withdraw and invest elsewhere. This cycle is then repeated indefinitely.

Proposition 1: Given credible public disclosure of initial contracts, an infinitely elastic supply of managers and the further assumption that

$$\delta > \frac{1}{1+q},$$

a *necessary* condition for equilibrium is that the initial managerial contracts must be such that managers are induced to act honestly and no managers are dismissed.

Proof: If $\Phi = 0$, managerial compensation is independent of the insiders' profits – so the manager has no incentive to cheat outsiders. An executive without a stake in the firm would not cheat outside investors directly. If contracts are credibly disclosed to the public, investors would shun any firm that offers its manager $\Phi > 0$. All firms accordingly are compelled to offer $\Phi = 0$[7]. The manager however could steal a maximum of $\bar{\varepsilon}$ from the board, though at the cost of possible exposure. This could be prevented if boards set the manager's salary at a sufficiently high level A^* and made the threat that following public exposure, the manager would be fired, and never rehired in the corporate sector again. This threat is credible because there is an infinitely elastic supply of managers. In this case, the manager will weigh his maximum gains from stealing against the expected value of his future salary losses in the event of exposure. The collective nature of the threat is important because if the fired executive expected to be hired again by other firms, the threat would lose its bite[8]. We need two conditions for this to work – the first ensuring that in the circumstances, honesty is incentive-compatible

[7]This is relevant to the initial contract announced by the boards. We will come later to the possibility of secret renegotiation.

[8]This does not require co-ordination among boards. Other boards – whose managers have not cheated – already have their incumbent managers and in any case also have access to an elastic supply of potential managers.

for the executive, and the second ensuring that boards find it worthwhile to suppress cheating even at the cost of paying a high salary. The first condition amounts to

$$\bar{\varepsilon} \leq \frac{\delta q}{1-\delta} A^* \tag{1}$$

Given that boards want to pay the lowest salary possible in the interests of profitability, they pay just enough to make inequality (1) an equality. So salary is set at

$$A^* = \frac{1-\delta}{\delta q} \bar{\varepsilon} \tag{2}$$

The second condition would be fulfilled if the high salary defined by (2) is still less than the amount the executive could have stolen. Then boards would have an incentive to pay this salary in order to rule out cheating. So we need

$$A^* \leq \bar{\varepsilon} \tag{3}$$

In combination with (2), this gives us our condition

$$\delta > \frac{1}{1+q} \tag{4}$$

So both conditions are fulfilled for a sufficiently high discount factor or a high enough probability of detection of cheating. We note that the threat of not employing an executive fired by another firm and exposed as a cheat by the public signal is essentially costless to other firms – given that managers are in infinitely elastic supply - and is therefore credible. Thus given (4) and public observability of contracts, we have a situation in which managers do not cheat either outside investors or their boards, and are not dismissed either. **Q. E. D.**

For the rest of our analysis we focus on the case where (4) is satisfied.

Boards would prefer the managers to cheat outside investors as long as

$$\frac{DSL}{H} > q\frac{\delta[(H-D)(F+S)-A^*]}{1-\delta}$$

$$\text{or} \quad q < \frac{(1-\delta)DSL}{\delta H[(H-D)(F+S)-A^*]}$$

The derivation of this is as follows. The one-time gain that would accrue to boards as a result of the manager's cheating is $\frac{DSL}{H} = \frac{DS(G-B)}{H}$ - the excess of the dues to outsiders payable in the good state over the outsiders' bad state dues. For cheating to be worthwhile, this should exceed the present discounted value of the losses to the board in the event of possible exposure by the public signal. Now the size of these losses, when assets are redeployable, is

$$H(F+S)-DS-\max\{D,R\}F-A^*$$

Here, the first term represents expected total profits in the industry. The second term shows what is paid to outsiders and is therefore subtracted from insiders' profits. The third term reflects the redeployability of assets : in the event of exposure and dissolution of the firm, insiders have the option of becoming outsiders in other firms, or of availing themselves of the "outside option" which pays R. This fact reduces the prospective losses from exposure. So does the fourth term because, if exposed, boards will no longer have to pay out the salary A^* to their managers. Now, since D will never fall below R (because of the outsiders' participation constraint), the prospective losses simplify to

$$(H-D)(F+S)-A^*.$$

Thus the condition that boards prefer cheating is :

$$\frac{DSL}{H} > q\,\frac{\delta[(H-D)(F+S)-A^*]}{1-\delta}$$

$$\text{or}\quad q < \frac{(1-\delta)DSL}{\delta H[(H-D)(F+S)-A^*]}$$

Were this the end of the story, the firm would be walled up in an 'honesty prison'. Given the initial contract, its executive, regardless of type, would have no incentive to cheat even though it wants him to do so. What is more, it cannot dismiss him for this act of defiance- such a dismissal, following an act of honesty (a payout of DG/H) would brand a firm as a potential cheat and precipitate an exodus of investors.

4. A Way out of Jail?

However, it is vital to consider the possibility of renegotiation. The board has a strong incentive to secretly renegotiate the contract it has publicly offered its executive.

If secret renegotiation is feasible, the firm may offer a secret contract (Φ_c, A_c), different from the (0, A^*) necessitated initially by mandatory disclosure or public observability of initial contracts. We show below that, though feasible, this is not an optimal strategy.

Proposition 2: A secret, renegotiated contract is worthwhile only if the ratio of outsider to insider capital s exceeds a certain threshold definable in terms of observable parameters, given by inequality (11") below.

Proof: Let Y_c be the total expected payoff of an executive who cheats in such a firm.

$$Y_c = \Phi_c[H(F+S)-DS + p\frac{DSL}{H}] + (1-\Phi_c)[A_c + \varepsilon] + \delta(1-q)Y_c$$

The first two terms on the right hand side give the manager's expected payoff from cheating. The cheating gains are multiplied by p because cheating in this model occurs conditionally on good luck. The last term indicates that cheating executives can continue to realize their discounted payoffs from cheating if and only if they are not caught by the public signal. Here we assume that executives caught cheating and fired are never rehired and have no outside option. Rehiring would mark firms out to the public as firms who wish to cheat.

We have

$$Y_c = \frac{\Phi_c[H(F+S)-DS+pDSL/H]+(1-\Phi_c)(A_c+\varepsilon)}{1-\delta(1-q)}$$

Executives prefer the renegotiated contract to the one publicly imposed in the honesty prison only if

$$Y_c \geq \frac{A*}{1-\delta} \tag{5}$$

If (5) is satisfied, an executive offered Y_c will refrain from cheating his board. Thus the by-product of a new contract acceptable to managers is that such a contract would also ensure that the firms seeking to renegotiate are not cheated by their own managers.

However, successful renegotiation also has two other requirements. First, one should be able to rule out the possibility of executives accepting the firm's renegotiated contract but then refusing to cheat. An executive who did this cannot be fired; payoffs and firings being common knowledge, such a dismissal would expose the firm as a potential cheat[9]. This requirement implies:

$$Y_c \geq \frac{\Phi_c[H(F+S)-DS]+(1-\Phi_c)A_c}{1-\delta} \tag{6}$$

Can we characterize the contract (Φ_c, A_c) more specifically? Since the expected income of the board in the event that it cheats is negatively related to the expected income Y_c that it must concede to its manager, the board decides the contract that it offers its manager by minimizing Y_c subject to the constraints (5) and (6). It minimizes

$$Y_c = \text{Max}\left[\frac{A^*}{1-\delta}, \frac{\Phi_c[H(F+S)-DS]+(1-\Phi_c)A_c}{1-\delta}\right] + \sigma \tag{7}$$

where σ is a minuscule differential between Y_c and the terms within square brackets.

This exercise implies

$$A^* = \Phi_c\{H(F+S)-DS\}+(1-\Phi_c)A_c. \tag{8}$$

The firm can offer its executive any contract that satisfies (8). In that event,

$$Y_c = \frac{A^*}{1-\delta} + \sigma \tag{9}$$

A second requirement is that boards should themselves have an incentive to renegotiate rather than stick to the initial contract. This implies

$$\frac{H(F+S)-DS+pDSL/H}{1-\delta(1-q)} - Y_c \geq \frac{H(F+S)-DS-A^*}{1-\delta}$$

or

$$Y_c \leq \frac{A^*}{1-\delta} + \frac{\{(1-\delta)pDSL/H\}-\delta q\{H(F+S)-DS\}}{(1-\delta)(1-\delta(1-q))} \tag{10}$$

A necessary and sufficient condition for (10) to be satisfied simultaneously with (9) is

$$\frac{pDSL}{H} - \frac{\delta q\{H(F+S)-DS\}}{1-\delta} > 0 \tag{11}$$

---

[9] Note that an instance of an executive deliberately not cheating when there is an opportunity to do so implies in our model that he pays outsiders their "good luck dues" instead of their "bad luck dues" when luck is good. Thus if outsiders observe that a manager is fired after they have got a good payoff, they can immediately infer that the board intends to cheat.

The expected gains from one-time cheating should exceed the present value of the stream of expected losses from the next period onward in the event of exposure. (11) implies

$$q < \frac{ps(1-\delta)LD}{\delta H[H(1+s)-Ds]} \qquad (11').$$

Here we use s to denote the ratio of outsider to insider capital. Alternatively, (11) can be written as a lower limit on s, the ratio of outsider to insider capital:

$$s > \frac{\delta q H^2}{(1-\delta)pDL - \delta q H(H-D)} = s^h(D) \qquad (11'')$$

Q.E.D

5. Back to Prison

The threshold $s^h(D)$ represents a credibility ceiling for the outsider-to-insider capital ratio. At any $s > s^h(D)$, firms, though they may announce a managerial contract that constrains their managers to honest behavior, will renegotiate with the latter a secret contract that induces them to cheat investors. Outsiders, knowing this, will avoid investing unless s is reduced to $s^h(D)$ or lower. All firms, on the other hand, will drive s up to $s^h(D)$, since, for any given D < H, their profits are an increasing function of s. On each unit of outsider capital they attract, they make a profit of H-D. Moreover, D will never exceed H as this would violate the insiders' participation constraint, since in that case being an outsider would clearly be more attractive than becoming an insider. The inequality (11") reduces therefore to an equation that represents essentially the firm's demand for outside capital as a function of the market return D on outside share capital. Thus firm size is pinned down.

This is subject however to the firm's participation constraint - D must never exceed the rate of return on entrepreneurial capital since that would prompt the entrepreneur to close down his business and invest in other firms instead. The implication is

$$H(F+S) - DS - A^* \geq DF$$

$$\text{or } H - \frac{A^*}{F+S} \geq D. \qquad (12)$$

Interestingly, we see from constraint (12) that firms with larger insider equity (and therefore more total equity – given that (11'') becomes an equality) find it easier to meet the participation constraint. This is an instance of a "vicious circle" in effect, a vicious circle springing from the fact that poorer firms cannot pay the manager his "efficiency wages"(wages high enough to ensure honesty) and cannot therefore start an enterprise. Thus if (12) is not met, potential entrepreneurs find outside investment more profitable.

All this corresponds to a game-theoretic equilibrium in which (a) investors believe that any firm that exceeds the limit $s^h(D)$ will secretly renegotiate a contract with its manager that induces him to cheat and so withhold their funds from such firms, (b) firms know this and expand up to this limit, but never beyond it, (c) there is therefore no secret contracting and no cheating in equilibrium. This helps us in pinning down firm size explicitly so that renegotiation, though feasible, does not take place in equilibrium.

One of the two constraints defining a floor on personal wealth required to enter entrepreneurship is derived from (12). It is easy to see that (12) can be rewritten as

$$F(1+s^h(D)) > \frac{A^*}{H-D} \qquad \text{(using 11'')}$$

or

$$F > \frac{A^*}{(H-D)(1+s^h(D))} = F^* \qquad (13)$$

While constraint (13) defines a floor on entrepreneurial wealth that springs from the entrepreneur's need to pay the manager high enough wages for honesty to be credible, there is a second floor which arises from the minimum size requirement for an enterprise to be operational. Again, using (11''), for the enterprise to come into being, we require

$$\frac{\text{I-F}}{\text{F}} \leq \text{s}^\text{h}(D)$$

or

$$F > \frac{\text{I}}{1+\text{s}^\text{h}(D)} = F^{**} \qquad (14)$$

Both the floors F^* and F^{**} are decreasing in s^h and increasing in D. The intuition is that a high D implies a high payout to outside investors, increasing the insiders' temptation to cheat via renegotiation. To offset this, the ratio of outsider to insider capital has to be low. This in turn implies that only those with high personal wealth can enter entrepreneurship, given the difficulty of mobilizing much outsider capital.

We now combine our game theoretic analysis with simple general equilibrium techniques to endogenize D and characterize equilibria for different possible parameter values. We will show that equilibrium is unique for a given initial wealth distribution.

Now which of the constraints (13) or (14) is binding depends on parameter values. First we consider the case where (13) is the binding constraint, or $F^* > F^{**}$. This happens when

$$H - \frac{A^*}{\underline{I}} < D$$

for which a sufficient condition is

$$H - \frac{A^*}{\underline{I}} < R \qquad (15)$$

given the outsiders' participation constraint,

$$D \geq R \qquad (16)$$

In this case, a person whose wealth satisfies (13) also automatically satisfies (14), so $F^*(s^h(D))$ becomes the relevant floor on entrepreneurial wealth.

Now suppose K is the aggregate wealth of the economy and P(W) the fraction of total wealth owned by those with wealth below W. Then the total demand for outside capital generated by the entrepreneurs who can enter is

$$X_d = K[1 - P\{F^*(s^h(D))\}]\, s^h(D) \qquad (17)$$

Here, the term in square brackets represents the ratio of entrepreneurial capital to the total wealth of the economy. The RHS, therefore, represents the amount of outside capital that entrepreneurs can apply for without compromising their credibility. As D falls, not only can each firm credibly apply for more outside investment, but also more firms can enter and create additional demand for capital (as F^* also falls). We can easily check that the demand for outsider capital is zero at D = H and increases steadily as D falls.

The total supply of outside capital is the total wealth of those below the minimum threshold required for entry:

$$X_s = KP[F^*(s^h(D))] \qquad (18)$$

This is subject to (16), the outside investors' participation constraint. The supply curve of outsider capital thus has a horizontal stretch at D = R and then slopes upward. Given the demand and supply curves defined in (17) and (18), there can be two kinds of equilibrium, depending on the initial wealth distribution :

(1) An interior equilibrium with H > D > R (Fig 1). The equilibrium D solves

$$s^h(D) = \frac{\mathrm{P}[\mathrm{F}^*(s^h(D))]}{1-\mathrm{P}[\mathrm{F}^*(s^h(D))]} \tag{19}$$

This is unique given the monotonicity of supply and demand. The interior equilibrium obtains for parameter ranges where

$$P(F^*(s^h(R)) < s^h(R)[1-P(F^*(s^h(R))] \tag{20}$$

In this equilibrium, outside capital is fully employed in industry, with none of it in the outside option. Therefore the return D on outsider capital strictly exceeds R. The optimal ratio of outsider to entrepreneurial capital just matches the ratio of wealth owned by those below the minimum wealth requirement for entry to that owned by those above and is uniquely determined by the distribution of wealth.

(2) An "excess supply" equilibrium in which D = R, the participation constraint of the investor binds and investors are indifferent between investing in the firm and outside[10] (Fig. 2). This equilibrium occurs for parameter ranges such that

$$P(F^*(s^h(R)) > s^h(R)[1-P(F^*(s^h(R))]. \tag{21}$$

There also remains a possibility that no market may exist. If W_{max} is the wealth of wealthiest individual, and $D_{\min}$ solves

$$s^h(D_{\min}) = \frac{A^*}{(H-D_{\min})W_{\max}} - 1$$

then, for all $D \geq D_{min}$, $X_d = 0$. Demand for outside capital is positive only for $D < D_{min}$. Now, if $D_{min} < R$, the minimum supply price of outside capital, the demand and supply curves will not intersect. No equilibrium will be possible (Fig. 3). Note that a low W_{max} implies a higher $s^h(D_{\min})$, and therefore a low D_{min}. Thus in a uniformly poor society, a market is unlikely to exist.

[10] However, investors will take care that their investment in the firm is not so much that it tempts cheating.

What if the binding constraint on entrepreneurial wealth were (14), not (13)? This would happen if

$$H - \frac{A^*}{\underline{I}} > D$$

In this case, F^{**} replaces F^* as the relevant threshold in (17)-(21). The derivation of demand and supply curves is similar. However, now the demand curve will have a horizontal stretch at $D = H - \frac{A^*}{\underline{I}}$ and will begin to decline thereafter (D cannot exceed this limit). The presence of the horizontal stretch implies that, in addition to the interior equilibrium and the excess supply equilibrium, a third kind of equilibrium may be possible for certain configurations of the initial wealth distribution. This is the equilibrium where equilibrium occurs along the horizontal stretch of the demand curve. In this equilibrium, $D = H - \frac{A^*}{\underline{I}}$ so that both the floors F^* and F^{**} become the same. This happens when

$$s^h(H - \frac{A^*}{\underline{I}})[1 - P(F^{**}(s^h(H - \frac{A^*}{\underline{I}}))] > P(F^{**}(s^h(H - \frac{A^*}{\underline{I}}))) \qquad (22)$$

Again, it is possible that a market may not exist, if no one is sufficiently wealthy. In this case, this would happen if $D_{min} < R$ where D_{min} solves

$$s^h(D_{\min}) = \frac{I}{W_{\max}} - 1 .$$

As already emphasized, for a given range of parameters, there is only one equilibrium (and for some, there are none). High δ, q or H or low L lead to a higher s^h for a given value of D. Thus if agents care more about the future, if public transparency is high or if the expected profit in the industry is high, relative to the cheating gains, a greater fraction of outsider capital can credibly be demanded for the same payout rate D. These factors would make the "excess supply" equilibrium less likely to obtain compared to the other kinds of equilibrium. When (13) binds – which happens for relatively low $\underline{I}$ - relatively high values of A* (which in turn implies low δ, low q or high $\bar{\varepsilon}$) can lower the equilibrium D, making likelier either an "excess supply"equilibrium or no equilibrium at all (and therefore no industry). If instead (14) binds, a high value of $\underline{I}$ lowers the equilibrium D and has a similar effect. The intuition is simply that large indivisibilities imply that a large amount of capital is necessary to set up enterprise. To be compatible with credibility, this large ratio of outsider capital has to be balanced by a low D. If the level of D compatible with credibility is below R, however, there will be no market.

5. Discussion

Our model has two different sets of implications. One set relates to insights about firm size while the other relates to developmental implications.

Outside investors in our model would recognize that even if the law mandated or competition compelled public observability and transparency of managerial contracts, the possibility of false disclosure and secret renegotiation would persist. Interestingly, it is the possibility of renegotiation that effectively limits the size of the firm – outsiders recognize that if their financing exceeds a certain threshold, boards' and managers' incentives to renegotiate and cheat may become too powerful.

All this bears on an age-old controversy in economics regarding the long-run cost curve and the limits to firm size in long-run equilibrium. While the rising portion of the U in short run cost curves can easily be attributed to fixed factors, factors are not fixed in the long run. Therefore the U shape of long run cost curves, in particular its upward-rising segment, has sometimes been attributed to increasing pressure on the limited coordinating capacities of the entrepreneur (Kaldor,1934). Coordination is a function that is unitary by its very definition because if the number of coordinators multiplies, the problem arises of coordinating the coordinators. Therefore, as the scale of production increases, the strain on the coordinator grows, resulting in inefficiencies of control and coordination. However, equilibrium implies by definition that all coordination problems have been resolved – so that this particular limit on the size of the firm cannot exist in long run equilibrium[11]. Quoting Kaldor (1934),

"For the function which lends uniqueness and determinateness to the firm – the ability to adjust, to co-ordinate – is an *essentially dynamic function* ; it is only required so long as

[11] This point closely follows Kaldor (1934).

adjustments are required ; and the extent to which it is required … depends on the frequency and the magnitude of the adjustments to be undertaken. It is essentially a feature not of "equilibrium" but of "disequilibrium"; it is needed only so long as, and in so far as, the actual situation in which the firm finds itself deviates from the equilibrium situation. With every successive adjustment to a given constellation of data, the number of "co-ordinating" tasks still remaining becomes less and the "volume of business" which a given unit of co-ordinating ability can most successfully manage becomes greater; until finally, in a full-period long-run equilibrium (in Marshall's stationary state), the task of management is reduced to pure "supervision," "co-ordinating ability" becomes a free good and the technically optimum size of the individual firm becomes infinite (or indeterminate)…"

Kaldor thus disagreed with the views of other scholars that increasing costs of co-ordination could limit the size of the firm in long-run equilibrium. However, there was a consensus that to pin down firm size, what was needed was a factor whose supply was limited at the level of the *firm,* even in the long run, but which had a definite supply price as far as the *industry* was concerned.

In our model, the size of the firm is limited by credibility concerns. Given asymmetric information (with regard to firm performance), outside investors realize that if they provide more than a certain amount of financing to any one firm, that firm's temptations to cheat them may be too strong. This generates a scarcity of an essential input – outside capital – in the *long* run. Outside capital is limited at the level of the firm, but has a definite supply price D for the industry as a whole – a supply price which we have endogenized by embedding our game theoretic model in a simple general equilibrium framework. Moreover, the reason why outside capital is limited, at the level of the firm – becoming perfectly inelastic beyond a point – lies in moral hazard and not in any technological properties of production or cost functions. Whether or not we agree with Kaldor's views on the inability of the traditional framework to result in a determinate firm size in long-run equilibrium, our work highlights a novel factor which helps to avoid such indeterminacy.

We now turn to some developmental implications of our model[12]. In poor countries, industrialization is hampered not just by an aggregate scarcity of wealth but by difficulties in mobilizing and concentrating it to support large-scale industry if it is too thinly spread. This is due in part to the well-known constraints on borrowing. Our model demonstrates that raising share-capital is also subject to a credibility ceiling, a limit on the ratio of outsider to insider capital set by concerns over cheating. This is one reason why share markets are underdeveloped in most poor countries and why firms in early modern Japan, Korea and India and within the Chinese Diaspora relied so heavily on extended family groups within which credibility was less of a concern than in an anonymous share market. This is also perhaps the reason why governments like the Korean deliberately skewed income distribution towards the chaebol, enabling the accumulation of personal fortunes that could help in building up credible large-scale industries[13].

If a market exists, if, for instance, we have a regular interior equilibrium, the return to outside capital D will decrease as the distribution of a given aggregate wealth becomes more equitable: if P( ) is higher for any W, the equilibrium ratio of outsider to insider capital s^h will be higher, the demand curve for outside capital $X_d(D)$ will shift leftward, the supply curve $X_s(D)$ will shift rightward, so that D falls. With increasing equity, the equilibrium level of D may fall below R, so that the market disappears.

This view of equity as inimical to industrialization contrasts strongly with received doctrine. Murphy, Shleifer and Vishny [1989], for instance, see equity as the basis of a homogeneous mass market for manufactures that fosters industrialization. This, however, is a demand-side phenomenon. It affects output only if the production pattern reflects the consumption pattern, as it must in a closed economy. In a small open economy, the two are independent.

Given the same Lorenz curve, a higher aggregate wealth facilitates industrialization. P(W) is smaller for a given level of P, so that, other things being equal, outside investors will enjoy a higher expected income D. A wealthier economy finds it easier to sustain a credible capital market. We have yet another factor that tends to make industrialization a cumulative process and yet another vicious circle of poverty.

---

[12] A related paper, in which however the agency problem within the firm is assumed away, is Guha (2005).
[13] Lal and Myint (1996) provide a good discussion of this.

Finally, if the minimum-size constraint (14) binds, rather than the "honesty wage" constraint (13), an increase in minimum firm-size $\underline{I}$ with the same level and distribution of wealth will increase P for any given s*: the demand curve for outside capital will fall and the supply curve rise, reducing D and increasing s* in equilibrium. Technological indivisibilities make for missing markets. If, however, the "honesty wage" constraint binds, an inaccurate public signal (poor transparency) or excessive impatience would raise the wage consistent with credibility, and increase the likelihood of market failure.

If the aggregate ratio of outsider to insider capital can be concealed (for instance by secret selling of insider stake) cheating can occur, as the actual ratio might well be raised beyond the safe limit. Empirical evidence – such as the studies already mentioned of Joh (2003) and Lemmon and Lins (2003) suggests that a high ratio of outsider to insider capital intensifies moral hazard. It also suggests that cheating is negatively associated with the salary component of executive pay (Peng and Roell (2004)), and that the separation of managers from compensation committees has not lowered executive pay (Anderson and Bizjak (2003)). All these findings bear out our conclusion that paying the manager enough – even when he has no say in the matter – may be a condition for credibility, provided this compensation is independent of insider profits.

6. Conclusion

We embed a repeated game with imperfect information and stochastic market outcomes between boards, managers and outside shareholders in a general equilibrium setting. We show that a renegotiation-proof managerial contract that deters cheating is possible only if outsider stake in the firm does not exceed a credibility ceiling. The personal wealth of the entrepreneur thus limits the size of the firm, a limit that arises not from technology but from moral hazard. A novel twist is thus provided to the age-old controversy regarding long run equilibrium firm size.

Further, the personal wealth of the entrepreneur must be enough for his firm both to reach the minimum viable size and to pay the manager enough to induce honesty. Those with less wealth can only be outside shareholders. Thus, the wealth distribution uniquely determines the equilibrium where it exists (including the rate of return to capital). For certain configurations of the wealth distribution, the equilibrium (and therefore the industry) disappears.

Our model helps explain the proliferation of family firms in relatively poor countries, moral hazard induced vicious circles retarding the development of poorer economies relative to richer ones with the same degree of inequality, and points to the growth effects of creating wealth inequality in a poor but open economy. It also provides a rationale for paying managers substantially more than their opportunity cost – but only in the form of salary.

References:

Aghion, Philippe and Patrick Bolton (1992): "An Incomplete Contracts Approach to Financial Contracting", *Review of Economic Studies*, 77: 473-494

Anderson, Ronald C. and John M. Bizjak (2003): "An empirical examination of the role of the CEO and the compensation committee in structuring executive pay", *Journal of Banking and Finance*, 27: 1323-1348.

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

Bernanke, Ben and Mark Gertler (1989) : "Agency Costs, Net Worth and Business Fluctuations", *American Economic Review*, 79:14-31.

Crystal, Graef S. (1991): "In Search of Excess : the Overcompensation of the American Executive", *W.W Norton, New York*.

Diamond, Douglas W. (1984): "Financial Intermediation and Delegated Monitoring", *Review of Economic Studies*, 51: 393-414.

Dixit, Avinash (2003): "On Modes of Economic Governance", *Econometrica*, 71: 449-481.

Faure-Grimaud, Antoine and Denis Gromb (2001): "Public Trading and Private Incentives", *Review of Financial Studies*, 17(4):985-1014.

Gale, Douglas and Martin Hellwig (1985): "Incentive Compatibility Debt Contracts : the One Period Problem", *Review of Economic Studies*, 52: 647-663.

Galor, Oded and Joseph Zeira (1993) : "Income Distribution and Macroeconomics", *Review of Economic Studies*, 60: 35-52.

Greif, Avner (1991): "Contract Enforceability and Economic Institutions in Early Trade: the Maghribi Traders' Coalition", *American Economic Review*, 83(3) :525-48.

Guha, Brishti (2005) : "Honesty and Intermediation : Corporate Cheating, Auditor Involvement and Implications for Development", Princeton University, mimeo.

Holmstrom, Bengt and Jean Tirole (1993): "Market Liquidity and Performance Monitoring", *Journal of Political Economy*, 101 : 710-740.

Huffington, Arianna (2003) : "Pigs at the Trough : How Corporate Greed and Political Corruption are Undermining America", *Crown Publishers, NY.*

Innes, Robert D (1990): "Limited Liability and Incentive Contracting with Ex-Ante Action Choices", *Journal of Economic Theory*, 52: 45-67.

Jensen, M. (1986): "Agency Costs of Free Cash Flow, Corporate Finance and Takeovers", *American Economic Review*, 76: 323-329.

Joh, Sung Wook (2003) : "Corporate Governance and Firm Profitability : Evidence from Korea before the Crisis", *Journal of Financial Economics*, 68(2) : 287-322.

Kaldor, Nicholas (1934): "The Equilibrium of the Firm", *Economic Journal* , 44:60-76.

Milbourn, Todd T. (2003): "CEO reputation and stock-based compensation", *Journal of Financial Economics*, 68 : 233-262.

Lal, Deepak and Hla Myint (1996) : *The Political Economy of Poverty, Equity and Growth,* Clarendon Press, Oxford.

Lemmon, Michael and Karl Lins (2003) : "Ownership Structure, corporate governance and firm value : evidence from the East Asian financial crisis", *Journal of Finance*, 58(4) : 1445-1468.

Murphy, Kevin, Andrei Shleifer and Robert Vishny (1989) : "Income Distribution, Market Size and Industrialization", *Quarterly Journal of Economics*, 104(3):537-564.

Peng, Lin and Ailsa Roell (2004), "Executive pay, earnings manipulation and shareholder litigation", mimeo.

Robinson, Austin (1934) : "The Problem of Management and the Size of Firms", *Economic Journal*, 44: 242-257.

Shapiro, Carl and Joseph Stiglitz (1984) : "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, 74: 433-444.

Tirole, Jean (1996): "A Theory of Collective Reputations (with Application to the Persistence of Corruption and to Firm Quality)", *Review of Economic Studies*, 63:1-22.

Zwiebel, Jeffrey (1996): "Dynamic Capital Structure Under Managerial Entrenchment", *American Economic Review*, 86, 1197-1215.

Figure 1

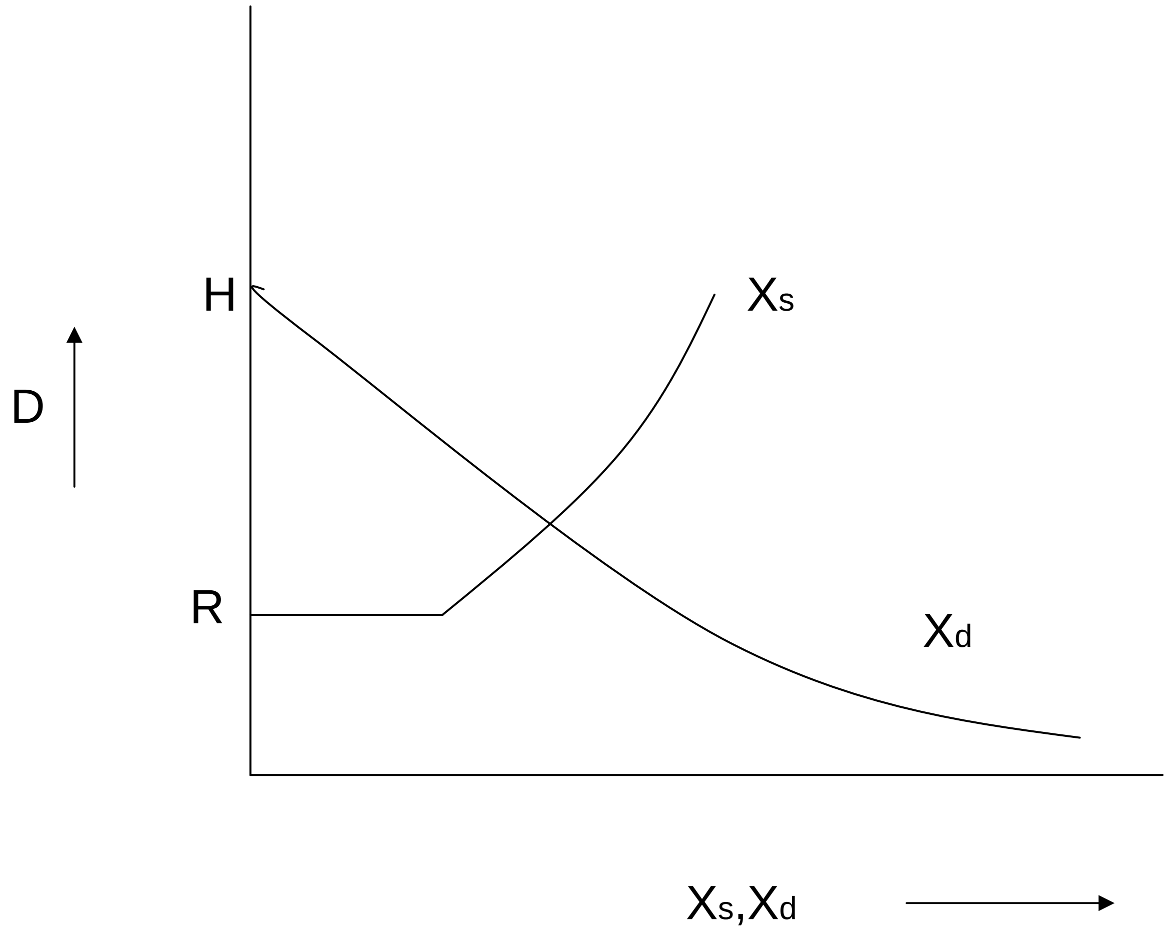

Figure 2

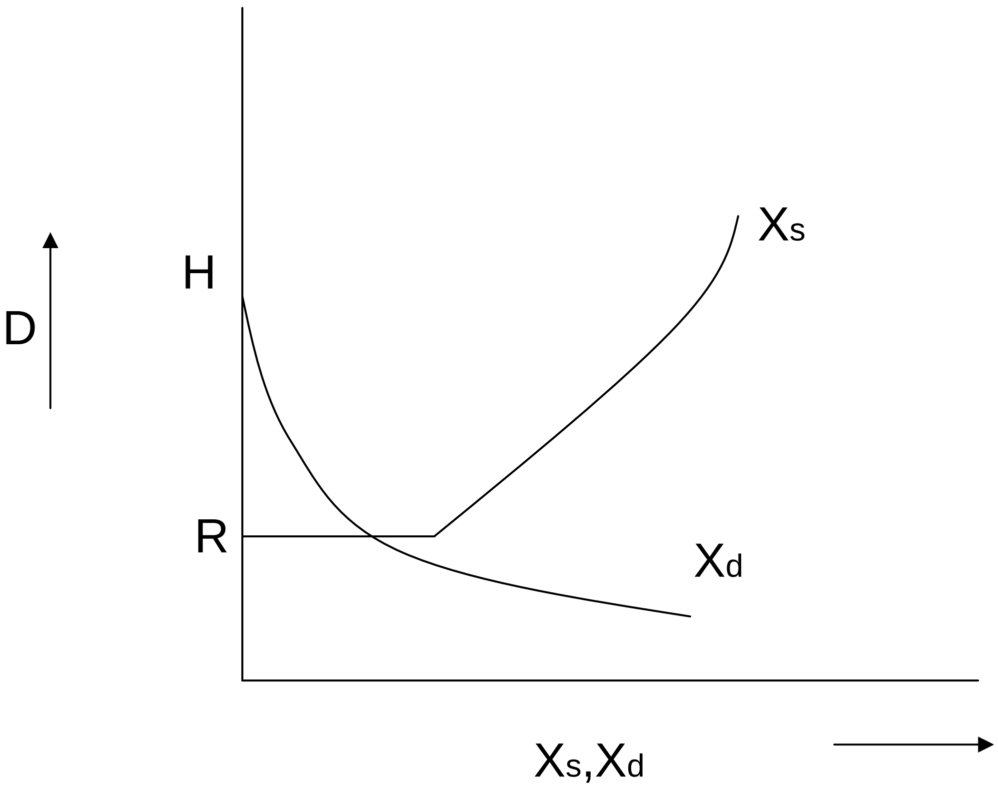

Figure 3