ESSAYS ON DIVIDEND POLICY

by

NALINAKSHA BHATTACHARYYA

B.Sc., North Eastern Hill University, Shillong, India, 1977
PGDM, Indian Institute of Management, Bangalore, India, 1986

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Commerce-Finance Division)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

May 2000

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of COMMERCE

The University of British Columbia
Vancouver, Canada

Date 27 Sep. 2000

DE-6 (2/88)

Abstract

This model explains dividend policy as a component of a screening contract set up by an uninformed principal. Our model assumes that the manager wants to maximize his net wealth and the principal recognizes this and sets up a screening contract to utilize the skill of the agent in the productive enterprise.This model deals simultaneously with *moral hazard* and *hidden information.*

We model *hidden information* in two ways-hidden information about the productivity of agent and hidden information about the initial cash endowment of the agent.

We find that when hidden information is about the productivity of the agent then, contrary to the findings of the dividend models based on the signaling paradigm, dividend - *conditional on cash availability* - bears an inverse relationship to managerial type. That is, for a given level of available cash, the lower type manager declares a higher dividend than that declared by a manager with higher productivity. Still this particular model can be used to explain many of the empirical findings obtained by other researchers. An interesting corollary of our model is that with costly effort and differences in productivity, the relationship between dividend and managerial type flips from being monotone increasing to monotone decreasing; this corollary extends the result obtained by Miller and Rock (1985). Another interesting implication of this dissertation is that dividends can be shown to be relevant in the presence of moral hazard and hidden information, even when the agency contract is chosen optimally .

When the hidden information is about the initial cash endowment of the agent then however we find that the agent with greater cash endowment declares higher dividend. We can therefore see that the nature of the hidden information plays a critical role in determining the relationship between dividends and agent type. When cash available is common knowledge and the asymmetric information is about the productivity of agent then higher dividend -conditioned on cash availability- is an indication of lower agent type. However when productivity of agent is common knowledge and the asymmetric information is about the initial cash endowment of the agent then higher dividend is an indication of higher agent type. We may note here that the outcome of the model in the second case is in line with the outcome of the signalling models and the free cash flow conjectures by Easterbrook (1984) and Jensen(1986).

Table of Contents

List of Tables

List of Figures

Acknowledgment

I have profited enormously from discussions with my supervisor, Prof. Ron Giammarino, and with the other members of my committee, viz. Prof. Jerry Feltham, Prof. Rob Heinkel and Prof. Burton Hollifield. My thanks to them. Prof. Vasant Naik and Prof. Raman Uppal have always been very generous with their time. Their genuine concern for my well-being helped smooth my progress through the Ph.D program. Prof. Amin Mawani and Prof. Mitch Farlee were always ready with supportive advice. My thanks to them. The changes suggested by the external examiner Prof. Praveen Kumar has improved the quality of this dissertation. My thanks to him. I also take this opportunity to thank the Canadian Commonwealth Scholarship program for supporting me during my Ph.D. Ruby Visser has been extra helpful in negotiating the maze of procedures and paperwork; my thanks to her. I would also like to record my thanks to Pat Darragh, Ken Wong, Eddie Baraghoush, Barbara Kent, Shaun Tanassee, Gary Schwartz, Khim Mah, Marjorie Stewart and Sally Bei for being very helpful and friendly, and for making the duration of the Ph.D very enjoyable. This dissertation is dedicated to my wife and son, without whose cooperation I would have bitten the dust long, long ago.

Dedication

This dissertation is dedicated to my wife Navneeta and son Aniruddha (Shubho) who have nurtured and supported me throughout the duration of the Ph.D program

Chapter 1

Introduction

Explaining dividend policy has been one of the most difficult challenges facing financial economists. Despite decades of study, we have yet to completely understand the factors that influence dividend policy and the manner in which these factors interact. Two decades ago, Black (1976) wrote, "The harder we look at the dividend picture, the more it seems like a puzzle, with pieces that just don't fit together" (p.5). The situation is pretty much the same today. In a recent survey of dividend policy, Allen and Michaely (1995) conclude that "[m]uch more empirical and theoretical research on the subject of dividends is required before a consensus can be reached" (p.833). The fact that a major textbook such as Brealey, Myers, and Marcus (1999) lists dividends as one of the ten important unsolved problems in finance reinforces this conclusion.

The seminal paper of Miller and Modigliani (1961) establishes that in a perfect capital market, given an investment policy, dividends are irrelevant in determining share value. Empirically, however, we have observed that a change in dividend policy does have a significant impact on the share price. Different researchers have concentrated on different types of imperfections in the market in order to understand the role of dividends. The two types of market imperfections that have been investigated are differential taxes on dividends and capital gains, and asymmetric information.

Considerations of differential taxes on dividend and capital gains have led to the *clientele theory* of dividend policy. The clientele theory says that shareholders face different tax rates with respect to dividends and capital gains. Shareholders wish to minimize their tax burden and as a result different shareholders have different levels of optimal dividend income. Shareholders sort themselves into clientele groups based on the established dividend policy of the firm in such a manner so that the individual shareholder has the optimum or near optimum dividend income. According to the clientele theory, if the dividend policy is changed for a firm, its shareholders reshuffle their holding so that their investment portfolios keep them in their desired level of dividend income.

Consideration of the second type of market imperfection, asymmetric information, has led to two classes of theories: the signaling theories, and the free cash flow hypothesis. The signaling theories posit dividend policy as a vehicle used by managers/insiders to transmit private information to the market. The free cash flow hypothesis, on the other hand, postulates that dividends are used to take-away excess cash from the mangers and put it in the hand of the shareholders.

In this thesis we explain dividend policy by using the asymmetric information paradigm. However, unlike the signaling models (where the *informed* manager/insider uses the dividend as a signaling device), we posit dividend policy as a component of a screening contract set up by an *uninformed* principal. In signaling models, *hidden information* is the source of informational asymmetry. In this work, we use a richer source of informational asymmetry — that due to

moral hazard (because the effort exerted by agent is not observable) and that due to *hidden information.* We assume that the manager wants to maximize his net wealth and the principal recognizes this and sets up a screening contract to utilize the skill of the agent in the productive enterprise.

We model *hidden information* in two ways-hidden information about the productivity of agent and hidden information about the initial cash endowment of the agent.

We find that when the hidden information is about the productivity of the agent then, contrary to the findings of the dividend models based on the signaling paradigm, dividend - *conditional on cash availability* - bears an inverse relationship to managerial type. That is, for a given level of available cash, the manager with lower productivity declares a higher dividend than that declared by a manager with higher productivity. We find that our model can be used to explain many of the empirical findings obtained by other researchers. An interesting corollary of our model is that when we include costly private effort and differences in productivity, the relationship between dividend and managerial type shifts from being monotone increasing to monotone decreasing. This relationship shows that incorporation of costly effort and difference in productivity modify the result Miller and Rock (1985) obtained. Miller and Rock study a model which does not include managerial effort. Another interesting implication of this dissertation is that dividends can be shown to be relevant in the presence of moral hazard and hidden information, even when the agency contract is optimally chosen.

Some of the empirical implications of this model (i.e the model where hidden information is about the productivity of the agent) are quite different from some of the implications of the signaling models and the implications of the free cash flow hypothesis. Under the signaling theories, higher firm value is signaled by higher dividends. Therefore, under the signaling paradigm, dividend increases should result in positive abnormal returns on the announcement of dividends. Also, *ceteris paribus*, the value of the firm should be an increasing function of the dividend. By contrast, our theory says that higher dividends, *conditioned on cash availability* (i.e., for a given level of cash availability; or, if earnings are taken as proxy for cash availability, then for a given level of earnings), is an indication of lower agent type and should result in a lower abnormal return and lower firm value (*ceteris paribus*). The free cash flow conjecture (Easterbrook, 1984; Jensen, 1986) posits that higher dividends are better because higher dividend removes free cash from the hands of the managers; consequently, the managers have less money to waste. According to this conjecture, announcement of higher dividends would also lead to higher abnormal return.

When the hidden information is about the initial cash endowment of the agent then however we find that agent with greater cash endowment declares higher dividend. We can therefore see that the nature of the hidden information plays a critical role in determining the relationship between dividends and agent type. When cash available is common knowledge and the asymmetric information is about the productivity of agent then higher dividend -conditioned on cash availability- is an indication of lower agent type. However when productivity of

agent is common knowledge and the asymmetric information is about the initial cash endowment of the agent then higher dividend is an indication of higher agent type. We may note here that the outcome of the model in the second case is in line with the outcome of the signalling models and the free cash flow conjectures by Jensen and Easterbrook.

The dissertation is organized as follows: Chapter 2 reviews the literature and motivates the dissertation. Chapter 3 discusses the basic two-agent model. In this model, there are two types of managers and the expected output of the firm depends on investment, managerial effort, and managerial type. Next, Chapter 4 deals with the case of a generalized production function and a continuum of agent types. Chapter 5 then discusses the pricing implications, and designs tests that can be used to empirically examine the refutable predictions of this dissertation. This chapter also discusses how far our theory is consistent with the existing empirical studies. Finally, Chapter 6 concludes the study. Proofs of propositions have been relegated, for the most part, to the Appendix.

Chapter 2

Literature Review and Motivation

2.1 The Stylistic Description of Dividend Policy.

The first empirical study of dividend policy was provided by Lintner (1956), who surveyed corporate managers to understand how they arrived at the dividend policy. Lintner finds that an existing dividend rate forms a bench mark for the management. Companies' management usually displayed a strong reluctance to reduce dividends. Lintner opined that managers usually have reasonably definitive target payout ratios. Over the years, dividends are increased slowly at a particular speed of adjustment, so that the actual payout ratio moves closer to the target payout ratio.

Bond and Mougoue (1991) reexamine the partial adjustment model of dividend payment suggested by Lintner. They find that when earnings follow a linear autoregressive process, then there are many combinations of target payout rate and the speed of adjustment that would fit the same earnings stream and dividend stream. They conclude that, for firms with autocorrelated earnings, Lintner's partial adjustment model gives results that are not unique; thus, for such firms the partial adjustment model is not a succinct description of dividend policy.

2.2 Dividend Irrelevance and Tax Clienteles

While Lintner (1956) provided the stylistic description of dividends, the watershed in the theoretical modelling of dividends was almost surely the classic

paper by Miller and Modigliani (1961), which first proposed dividend irrelevance. Essentially, their model is a one-period model under certainty. Given a firm's investment program, the dividend policy of the firm is irrelevant to the firm value, since a higher dividend would necessitate more sale of stock to raise finances for the investment program. The crucial assumption here is that the future market value will remain unaffected by current dividends.

The argument rests on the assumptions that the investment program is determined independently and that every stockholder earns the same return (i.e., the discount rate remains constant). Miller and Modigliani's (1961) dividend-irrelevance argument is elegant, but this does not explain why companies, the public, investment analysts are so interested in dividend announcements. Clearly, the observed interest in dividend announcement must be related to some violation of the Miller and Modigliani assumptions.

Miller and Modigliani, while formulating their famous dividend irrelevance propositions, observed that in the presence of taxation, investors will form clienteles with specific preferences for particular levels of dividend yields. This specific preference for dividends may be determined, *inter alia*, by the marginal tax rates faced by the investor. Altering the dividend level, according to Miller and Modigliani, leads only to a change in the clientele of shareholders for the firm.

Part of the dividend puzzle arises from the fact that dividends are typically taxed at a higher rate compared to the income from capital gains. We should, therefore, expect investors to prefer cash from capital gains over cash from dividends. Miller and Scholes (1978) provide an ingenious scheme to convert

dividend income to capital gains income. Their work provides a fresh rationale for the dividend irrelevance position. Their argument is based upon a common income tax provision which allows interest expenses to be deducted from income before applying tax. Miller and Scholes show that by borrowing an appropriate amount, the interest amount can be set off against the dividend income in a way that reduces the taxable income to zero. Miller and Scholes contend that the increases in risk due to borrowing can be countered by investing the borrowed amount in a risk-free insurance contract, where the amount accumulates at the risk-free rate. In this way, they argue, the tax shield on the interest expense can be used to neutralize the tax incidence on the dividend income without incurring any additional risk due to increased borrowing.

Peterson, Peterson, and Ang (1985) look at the extent to which investors attempt to shield their dividend income from taxation, a la Miller and Scholes (1978). They look at returns selected from individual income tax returns filed. They find that 85% of the filed returns report no dividend income. They also find that about 56% of those assesses reporting dividend income, do not take advantage of interest deduction schemes a la Miller and Scholes.

2.3 Informational Asymmetry and Signalling Models

Deviations from the Miller and Modigliani (1961) dividend irrelevance proposition is obtainable only when the assumptions underlying the setting of Miller and Modigliani are violated. The tax-clientele hypothesis uses the market imperfection of differential taxation of dividends and capital gains to explain the

dividend puzzle. Bhattacharyya (1979) develops another explanation for the dividend policy based on asymmetric information. The managers have private knowledge about the distributional support of the project cash flow and they signal this knowledge to the market through their choice of dividends. In the signalling equilibrium higher value of the support is signalled by higher dividend. In other words, the better the news, the higher is the dividend.

Heinkel (1978) considers a set up where different firms have different return-generating abilities. This information is transmitted to the market by means of dividends, or equivalently, from investing at less than the first best level. In the equilibrium of Heinkel's model, the firm with less productivity invests up to its first best level and declares no dividend, while the firm with higher productivity invests *less than its first best level of investment,* and declares the difference between the amount raised and the amount invested as the dividend. The firm with higher productivity acts in this way in order to distinguish itself from the firm with less productivity. Dividends are still irrelevant in the sense that both firm types could raise an extra X dollars with a new issue to pay an extra X dollars as a dividend with no signalling effect. The signalling cost in this model comes from reduced investment from first best level. In contrast, the signalling cost in Bhattacharyya (1979) comes from taxation and non-symmetric cost of raising funds in the capital market.

Bhattacharyya's and Heinkel's work was followed by a number of other papers which posited that dividends are used by managers to transmit information to the capital market. Notable works in signalling paradigm of

dividend policy are those of Miller and Rock (1985), John and Williams (1985), Ambarish, John, and Williams (1987) and Williams (1988). These signalling models typically characterize the informational asymmetry by bestowing the manager or the insider with information about some aspect of the future cash flow. In the signalling equilibriums obtained in these models, the higher the expected cash flow, the higher is the dividend. In Miller and Rock, the signalling cost is the opportunity cost of less than first best investment. In John and Williams; Ambarish, John, and Williams; and Williams, the differential taxation of dividends vis-a-vis capital gains sustains the signalling equilibriums. In these papers dividends sustain a fully separating equilibrium. By contrast, Kumar (1988) demonstrates that dividends could also sustain a semi-separating equilibrium where the manager has private information about the productivity of the firm.

Bar-Yosef and Venezia (1991) set up a rational equilibrium expectation model. Bayesian investors expect that dividends will be proportional to cash flows. Managers have advance *noisy* information about the future cash flow. The investors observe the dividend and update their belief about the cash flow. Under these circumstances, Bar-Yosef and Venezia show that the optimal dividend is proportional to the cash flow.

Brennan and Thakor (1990) focus on a different question compared to the other signalling type papers on dividend policy. Most dividend policy papers model the dividend decision, as a decision about the amount to be distributed as dividends. In contrast, this paper views the amount of cash to be distributed as

exogenously given. It considers three forms of disbursement: dividend declaration, non-proportionate share repurchase through open market operation, and non-proportionate share repurchase through tender offer. Brennan and Thakor assume that there are two classes of shareholders — informed and uninformed. They show that in a tender offer, the uninformed shareholder always tenders, whereas the informed holds onto his/her shares. The situation is reversed in an open market operation, where the informed shareholder always sell his/her holding and the uninformed never does.

2.4 Free Cash Flow Hypothesis

The rich theoretical development in modelling dividends as signals of private managerial/entrepreneurial information also gave rise to empirical research seeking to determine the fit of the signalling theory to real world data. Typically, the empirical literature[1] attempted to test the signalling paradigm counterpoised against an alternative rationale for dividends advanced by Jensen (1986), based on the Principal-Agent framework. According to this framework, dividends are used by shareholders as a device to reduce overinvestment by managers. The managers control the firm; therefore, they might invest cash in projects with negative net present values, but which increase the personal utility of the managers in some way. A dividend reduces this free cash flow and thus reduces the scope for overinvestment. The two most cited works in this genre are the papers by Easterbrook (1984) and by Jensen (1986). Unfortunately, neither

[1]For example see Lang and Litzenberger (1989), Denis, Denis, and Sarin (1994), Bernheim and Wantz (1995), Yoon and Starks(1995).

of these papers try to model the situation; rather, they put forward plausible hypotheses.

On the one hand, Easterbrook (1984) hypothesizes that dividends are used to take away the free cash from the control of the managers and pay it off to shareholders. This ensures that the managers will have to approach the capital market in order to meet the funding needs for new projects. The need to approach the capital markets imposes a discipline on the managers, and thus reduces the cost of monitoring the managers. Additionally, Easterbrook hypothesizes that the imperative to approach the capital market also acts as a counterweight to the managers' own risk aversion.

Jensen (1986), on the other hand, contends that in corporations with large cash flows, managers will have a tendency to invest in low return projects. According to Jensen, debt counters this by taking away the free cash flow. Jensen contends that takeovers and mergers take place when either the acquirer has a large quantum of free cash flow or the acquired has a large free cash flow which has not been paid out to stakeholders. Although Jensen does not deal with the issue of dividends, empirical researchers of dividend policy often use Jensen's article for motivating tests of the free cash flow hypothesis of dividend policy.

The empirical evidence on the three hypotheses are mixed, as we observe in **Table I** . Dividend policy thus continues to remain a puzzle. We can however enumerate some interesting stylized facts. In **Table II** we compile the stylized facts as they emerge from a study of the empirical literature.

Research	Dividend Clientele Hypothesis		Signalling Hypothesis		Agency Hypothesis	
	Do Not Reject	Reject	Do Not Reject	Reject	Do Not Reject	Reject
Aharony and Swary (1980)			✓			
Bernheim and Wantz (1995)			✓			✓
Black and Scholes (1974)		✓				
Chaplinsky and Seyhun (1990)	✓					
Chen, Grundy, and Stambaugh (1990)		✓				
Christie (1994)				✓		✓
Denis, Denis, and Sarin (1994)	✓		✓			✓

Table I: Findings of Empirical Research vis-a-vis the Three Hypotheses of Dividend Policy

Research	Dividend Clientele Hypothesis		Signalling Hypothesis		Agency Hypothesis	
	Do Not Reject	Reject	Do Not Reject	Reject	Do Not Reject	Reject
Dhillon and Johnson (1994)					✓	
Downes and Heinkel (1982)				✓		
Kao and Wu (1994)			✓			
Lakonishok and Vermaelen (1986)	✓					
Lang and Litzenberger (1989)				✓	✓	
Lewellen, Stanley, Lease and Schlarbaum . (1978)		✓				
Litzenberger and Ramaswamy (1982)	✓					

Table I: Findings of Empirical Research vis-a-vis the Three Hypotheses of Dividend Policy (Continued)

Research	Dividend Clientele Hypothesis		Signalling Hypothesis		Agency Hypothesis	
	Do Not Reject	Reject	Do Not Reject	Reject	Do Not Reject	Reject
Long, Malitz, and Sefcik (1994)						✓
Manuel, Brooks, and Schadler (1993)			✓			
Michaely (1991)		✓				
Penman (1983)			✓			
Poterba (1986)		✓				
Sant and Cowan (1994)			✓			
Smith and Watts (1992)					✓	
Yoon and Starks (1995)						✓

Table I: Findings of Empirical Research vis-a-vis the Three Hypotheses of Dividend Policy (Continued)

Sl. No	Stylized Facts	Source of Evidence and Remarks
1.	At least 25% of firms listed in NYSE do not pay any dividends.	"In fact, more than 25% of firms listed on the New York Stock Exchange do not pay any dividends at all"(Lee, 1996, p.33)
2.	Managers are very unwilling to reduce dividends.	Lintner (1956) noted this trend in his paper. DeAngelo and DeAngelo (1990) find that for 80 NYSE firms in financial crisis, managers are more willing to cut the level of dividend than to omit the dividend altogether.
3.	Large dividend reductions are valued more severely by the market than dividend omission	Christie (1994) finds that for less than 20% reduction, prices fall by about 4.95%, while for reductions exceeding 60%, prices fall by about 8.78%. In contrast, prices fall by about 6.94% for dividend omissions
4.	Dividend increases (decreases) are associated with increase (decrease) in capital expenditures in subsequent years.	Yoon and Starks (1995)

Table II: Stylized Facts on Dividends

Sl. No	Stylized Facts	Source of Evidence and Remarks
5.	For IPOs, the earnings multiple is less for dividend paying stock than for non-dividend paying stock. The value of the firm is equal to the product of earnings multiple and earnings.	Downes and Heinkel (1982) set up a regression linking the value of the firm with the product of earnings and a multiple. The multiple is assumed to be a function of the firm's many descriptors, one of which is a dummy variable which captures the information whether a dividend is paid or not. The sign of the coefficient of this variable is negative.
6.	With changes in dividends, price reactions for stocks are in the opposite direction to the price reactions for bonds.	Dhillon and Johnson (1994)
7.	Volatility of earnings, volatility of analysts forecasts and beta increase after dividend omission.	Sant and Cowan (1994)

Table II: Stylized Facts on Dividends (Continued)

Sl. No	Stylized Facts	Source of Evidence and Remarks
8.	The longer the company has been paying dividends the stronger is the reluctance of the managers to reduce dividends.	DeAngelo and DeAngelo (1990)
9.	Industries with high growth options pay less dividends.	Smith and Watts (1992) found this result with industry level data. Gaver and Gaver (1993) verify the results of Smith and Watts (1992) using a more rigorous methodology, and firm level data. They find that growth firms have lower debt/equity ratio and significantly lower dividend yields than non-growth firms. The growth firms also pay higher compensation to the executives than non-growth firms.
10.	Loss is a necessary condition for dividend reductions, but is not a sufficient one.	DeAngelo, DeAngelo and Skinner (1992)

Table II: Stylized Facts on Dividends (Continued)

Sl. No	Stylized Facts	Source of Evidence and Remarks
11.	Dividend reductions adversely affects the chances of managers obtaining outside directorships.	Kaplan and Reishus (1990)
12.	Investors value a dollar of cash dividend more than a dollar of stock dividend.	Poterba (1986) reached this conclusion with his study of the return behavior on two classes of stock issued by The Citizen's Utility — one class paid stock dividends, and the other class paid cash dividend.
13.	For large publicly traded firms, executive compensation is influenced by dividends.	Ferreira White (1996) examined the compensation contracts of 62 large companies in oil and gas, food processing, and defense/aerospace industries. Twenty-eight of these had a dividend provision. Pavlik, Scott, and Tiessen (1993) report that in 1991, base salary was 33% of executive compensation, while stock options, restricted stock and performance shares was 36%. This component is affected by dividend decision.

Table II: Stylized Facts on Dividends (Continued)

2.5 Dividend Policy and Executive Compensation

We would like to call attention to the last stylized fact contained in **Table II**, which suggests that executive compensation is influenced by dividends, either directly or indirectly, through the linkage of dividends to stock price. Ferreira White (1996) examined the compensation contracts of 62 large companies in oil and gas, food processing and defence/aerospace industries. Twenty-eight of these had a dividend provision. Pavlik, Scott, and Tiessen (1993) report that in 1991, base salary was 33% of executive compensation, while stock options, restricted stock, and performance shares were 36%. These components are affected by dividend decisions.

There are many ways to directly link dividends to managerial compensation. Dividend units have the most direct linkage as Larcker (1983) explains, "A *Dividend Unit* assigns a hypothetical fixed number of stock to the executive with the amount of compensation paid (either in stock or cash) being equal to the number of shares multiplied by the dividends per share paid to shareholders" (p.6). Restricted stock provides another mechanism to link dividends to compensation. Crystal (1989) writes "[With restricted stock] [t]he company simply gives the executive the shares with the restriction that he may not sell them for a specified period, generally five years. In the meantime, he gets the dividends and can vote the shares"(p.104). Crystal also lists the companies that use restricted stock. He examined 161 companies selected from Fortune 500 companies. Sixty-one of these 161 companies issued restricted stock to their CEOs.

We also have evidence that short term compensation is linked to dividends. Healy (1985) mentions that often the upper limit on the amounts to be transferred to a bonus pool is related to cash dividend payment on common stock. Consistent with Healy's work, Lewellen, Loderer, and Martin (1987) find a statistically significant positive relationship between the short term component of executive compensation (i.e., salary and bonus) and dividend payout. They write "[t]he findings further support the notion that firms seek to prevent tendencies toward over-retention of earnings by linking salary and current bonus payments to dividend payouts" (p.301).

The other important determinant of executive compensation is accounting earnings (Pavlik, Scott, and Tiessen, 1993; Healy, 1985). For example, Healy finds that "[b]onus contracts have a similar format to performance contracts except that they specify annual rather than long-term earnings goals" (p.87). Pavlik, Scott, and Tiessen state that "[a]ccounting and stock returns are the performance measures that have been consistently associated with changes in compensation"(p.155).

2.6 Motivation

So far we have just been underscoring the empirical observation that executive compensation is, in many cases, linked to stock prices, dividends, and accounting earnings. We need to understand these linkages, and the underlying economic rationales that motivate them. Our primary concern, however, is to understand dividends. We have seen earlier that the empirical evidence on the

major paradigms of dividend theory is mixed. To our minds, this mixed evidence indicates that a substantial gap exists in our understanding of why dividends are paid (or not paid, as in the case of about 25% of firms listed in NYSE). The linkage of executive compensation to dividends (either directly or through stock prices), accounting earnings, and stock prices also needs to be understood. Note that linking executive compensation to stock prices is feasible only for companies whose stocks are listed in stock exchanges. Executive compensation is typically determined by solution to agency issues. To quote Jensen and Murphy (1990)

> "The conflict of interest between shareholders of a publicly owned corporation and the corporation's chief executive officer (CEO) is a classic example of a principal-agent problem. If shareholders had complete information regarding the CEO's activities and the firm's investment opportunities, they could design a contract specifying and enforcing the managerial action to be taken in each state of the world. Managerial actions and investment opportunities are not, however, perfectly observable by the shareholders; indeed, shareholders do not often know what actions the CEO *can* take or which of these actions will increase shareholder wealth. In these situations, agency theory predicts that compensation policy will be designed to give the manager incentives to select and implement actions that increase shareholder wealth."(pp.225-226)

In this dissertation, we apply the principal-agent paradigm to understand the role of dividends as a control device. Given a particular level of cash flow,

shareholders will want this to be apportioned between dividends and further investment in the activity of the firm. However, the actual apportionment of this cash flow into investment and dividend is decided by the manager. The productivity of the manager (the "type" of manager) is private knowledge and so is the effort exerted by the manager. We investigate dividend/investment in this setting using the principal-agent paradigm. Although the two papers most cited by empirical researches on dividend policy in an agency paradigm are by Easterbrook (1984) and by Jensen (1986), neither of these papers formally models the agency relation. This dissertation will contribute to explanations of dividend policy by rigorously modelling the dividend decision in an agency theoretic setting.

Chapter 3

The Basic Two Type Agent Problem

3.1 Description of the Model

Our model contains one principal and one agent; both are risk-neutral. The risk-neutrality of the principal is representative of shareholders who hold well diversified portfolios. The assumption of agent risk-neutrality does not permit the first best solution because (as we shall see later) the model involves pre-contract private information. Furthermore, the principal owns an investible resource C. C can be conceptualized as the earning or cash flow from the previous period.

There are three factors of production: skill, effort and investment. The agent supplies the first two of these factors of production — namely skill and effort — and decides on the level of investment, the third factor. First, the agent has a skill which is necessary to make the asset productive. This skill is represented by a parameter θ. For our purpose, we assume that it can only have two values, L(ow) and H(igh) and $\theta_H > \theta_L$. Before entering into the contract, the agent knows his/her level of skill, although he/she cannot alter it; the principal, however, is not aware of the agent's skill level.

In addition to having the necessary skill, the agent also needs to exert effort in order to make the asset productive. The effort exerted by the agent, represented by e, is not observable by the principal. The agent chooses the level of effort supplied and incurs a private cost of effort. The cost of effort is represented by $\frac{1}{2}Me^2$, where M is a positive number.

The agent also selects I, the amount to be invested. We assume that the parameter values are such that $I \leq C$. The excess of C over I is paid off as a dividend D. That is $D = C-I$. The subscript H and L are used to indicate the choice for a particular type. That is, D_H and I_H are the choices of the H type agent. Similarly D_L and I_L are the choices of the L type agent.

Note that this budget constraint (namely, that the sum of dividend and amount actually invested is equal to C) implicitly assumes that the agent cannot lend or borrow in the capital market. In our model compensation is a function of dividend and output. We want to understand how such a compensation function impacts on dividends when different agents have different productivity. In the real world, however, managers issue shares, borrow money and pay dividends. To take into account all the three actions simultaneously in an agency-theoretic setting we would require a much richer model-possibly a multi-period game-theoretic model. Our conjecture is that in such a set up the information content of the action would be based on the net financing (issue minus dividend) which, when added to initial cash, gives investment. The market would have to price the new shares based on what they learn from this net financing. Incentive conditions become more interesting as the manager considers both investment distortions that arise from the non linear production function and the pricing implications for the new issue that results from truth telling or mimicking. Such a model would be more intricate and more interesting but it is beyond the scope of this dissertation. In our model therefore we would consider the budget constraint

(namely, that the sum of dividend and amount actually invested is equal to C) to be binding.

Finally, the realized production $\tilde{Y}$ is observed by both principal and agent. The production function is given by $\tilde{Y}=\theta$ e ln(I) + $\tilde{\varepsilon}$ where $\tilde{\varepsilon}$ is random noise with zero mean. For tractability we assume that the compensation contract offered to the agent is linear and is of the form $\tilde{\omega}_j = \beta_0\left(\hat{j}\right) + \beta_D\left(\hat{j}\right)D_j + \beta_Y\left(\hat{j}\right)\tilde{Y}$ where $\hat{j}$ is the type reported by the agent. The sequence of the foregoing events is shown in **Table III**.

1	2	3	4	5
C is observed by all. Agent privately observes his/her own type.	Principal offers a menu of wage contracts	Agent declares his/her type and selects from menu.	Agent declares D, spends effort and invests the balance of C over D.	The final output is realised and observed by all. Principal receives the payoff and agent receives the wage.

Table III: The Sequence of Events

3.2 Notation

For ease of reference, the notations used are summarized below.

C Cash available for investment

I_j Investment level chosen by agent of type j, when agent truthfully declares type to be j.

$I_{\hat{j}j}$ Investment level chosen by agent of type j, when agent declares type to be $\hat{j}, j \neq \hat{j}$

D_j Dividend level Selected by an Agent of type j, when agent truthfully declares type to be j = $C - I_j$.

$D_{\hat{j}j}$ Dividend level chosen by agent of type j, when agent declares type to be $\hat{j}, j \neq \hat{j}$; $D_{\hat{j}j} = C - I_{\hat{j}j}$

θ_j Productivity parameter for agent of type j. For our purpose we assume that j can only have two values - L(ow) and H(igh) and $\theta_L < \theta_H$.

e_j Effort put in by the agent of type j, when agent truthfully declares type to be j.

$e_{\hat{j}j}$ Effort put in by agent of type j, when agent declares type to be $\hat{j}, j \neq \hat{j}$.

M A parameter of the cost of effort function.

p Probability that the agent is of type L

$\tilde{\omega}_{\hat{j}j}$ The wage function offered to an agent of type j who reports his/her type to be $\hat{j}$, assumed to be linear in D and $\tilde{Y}$.

Thus $\tilde{\omega}_{\hat{j}j} = \beta_0\left(\hat{j}\right) + \beta_D\left(\hat{j}\right)D_{\hat{j}j} + \beta_Y\left(\hat{j}\right)\tilde{Y}$

where

$\beta_0\left(\hat{j}\right)$ is the constant component of the wage

$\beta_D\left(\hat{j}\right)D_j$ is the component of the wage proportional to dividend.

$\beta_Y(\hat{j})\tilde{Y}$ is the component of wage proportional to output.

$\tilde{Y}$ is stochastic output.

We use a caret (^) to indicate parameters as reported by agent about its type. Thus, $\hat{j}$ represents the type reported by agent. We shall use the generalized revelation principle as enunciated in Myerson (1982) and Faynzilberg and Kumar (1997) so that in equilibrium the agent tells the truth i.e., $j=\hat{j}$.[2]

3.3 The Full Information Solution

The full information solution will be given by the solution to the following optimization problem:

$$\underset{D_j,\, e_j}{\text{Maximize}} \;\; D_j + \theta_j e_j ln\left(C - D_j\right) - \tfrac{1}{2} M e_j^2, \text{ for each } \; j = L, H$$

The above optimization problem is drawn up by assuming that all the information is known to everybody. In that scenario, the agent would be paid only the reservation wage (i.e, the wage below which the agent will not work). So there are no information rents. In addition, the investment and effort levels are efficient,

---

[2] We have restricted ourselves to the class of linear compensation contracts. Revelation principle can be applied here to simplify our analysis by using the enunciation of the principle in Rasmusen (1989). Rasmusen states the revelation principle as "(f)or every contract ...that leads to lying... there is a contract .. with the same outcome ... but no incentive for the agent to lie"(p.198). Rasmusen's enunciation shows that revelation principle can be applied to the class of linear compensation contracts.

i.e., they maximize the size of the pie. In our subsequent analysis, we assume that the efficient levels are interior solutions.

If the interior solution is optimal, then the first order conditions for optimality, for each j, are[3]

$$1 - \frac{\theta_j e_j}{C - D_j} = 0$$
$$\theta_j \ln(C - D_j) = M e_j \tag{1}$$

3.4 The Principal's Problem

Let us now consider the second best solution. The Principal's problem can now be represented as in Program 1.

Program 1

$$\underset{\substack{\beta_0(L),\beta_D(L),\beta_Y(L),\\ \beta_0(H),\beta_D(H),\beta_Y(H),\\ D_L,e_L,D_H,e_H}}{Maximize} \quad \Big[p\big[D_L + \theta_L e_L \ln(C - D_L) - \beta_0(L) - \beta_D(L) D_L$$
$$-\beta_Y(L)\theta_L e_L \ln(C - D_L)\big] + (1-p)\big[D_H + \theta_H e_H \ln(C - D_H)$$
$$-\beta_0(H) - \beta_D(H) D_H - \beta_Y(H)\theta_H e_H \ln(C - D_H)\big]\Big]$$
$$D_j \in [0, C],\ e_j \in [0, \infty),\ j = L, H$$

such that the following incentive constraints and participation constraints are satisfied.

---

[3] We can show that the following conditions must be satisfied for the existence of first best interior optimal solution. $M < \theta_j^2$ and $D_j < C$ - Exp(1), where Exp(1) is the exponential function.

Incentive Constraints

$$\left(j, D_j, e_j\right) \in \underset{\hat{j},D,e}{Arg\max}\left[\beta_0\left(\hat{j}\right) + \beta_D\left(\hat{j}\right)D + \beta_Y\left(\hat{j}\right)E\left[\tilde{Y}\middle|\theta_j, D, e\right] - \tfrac{1}{2}M\,e^2\right]$$

$$j = L, H \qquad \text{(IC)}$$

Participation Constraints

$$\beta_0\left(j\right) + \beta_D\left(j\right)D_j + \beta_Y\left(j\right)E\left[\tilde{Y}\middle|\theta_j, D_j, e_j\right] - \tfrac{1}{2}M\,e_j^2 \geq \omega_0 \quad , \; j = L, H \qquad \text{(PC)}$$

where ω_0 is the reservation wage.

The maximand in Program 1 is obtained by considering the expected net gain to the principal. The incentive constraint captures the fact that the agent would choose dividend D and effort e so as to maximize his/her own expected utility. The participation constraint (which is also known as the individual rationality constraint) takes account of the fact that the agent will not work below a certain minimum wage.

3.5 The Agent's Problem

In this section, we consider the agent's problem in choosing D and e given β_0, β_D and β_Y. Once the agent has chosen a contract from among those offered by the principal, the coefficients β_0, β_D and β_Y can be treated as given parameters; the agent then solves the following problem for a given β_0, β_D and β_Y, and for a given productivity parameter θ. Therefore, we omit the arguments j (used to indicate the type of the agent) and $\hat{j}$ (used to indicate the type reported by agent) in writing the expressions in this section.

The agent's problem is

$$\underset{D,e}{Maximize} \quad \beta_0 + \beta_D\, D + \beta_Y\, E\left[\tilde{Y}\,\middle|\,\theta, D, e\right] - \tfrac{1}{2} M\, e^2$$

We concentrate on the case where agent's problem has an internal solution. We note that the first order approach is valid here because the agent's maximand is a strictly concave function. The first order conditions would therefore determine the local as well as global maximum for the agent.

The first order conditions for optimality are

$$\begin{aligned} &\beta_D - \frac{\beta_Y \theta\, e}{C - D} = 0 \\ &\beta_Y \theta \ln(C - D) = M\, e \end{aligned} \tag{2}$$

3.6 Solution To The Principal's Problem

The principal's problem can be represented in Program 1A.

<u>Program 1A</u>

$$\begin{aligned} \underset{\substack{\beta_0(L),\beta_0(H),\\ \beta_D(L),\beta_D(H),\\ \beta_Y(L),\beta_Y(H)}}{Maximize} \quad & [p[\{1-\beta_D(L)\}D_L + \{1-\beta_Y(L)\}\theta\, e_L \ln(C - D_L) - \beta_0(L)] \\ & +(1-p)[\{1-\beta_D(H)\}D_H + \{1-\beta_Y(H)\}\theta\, e_H \ln(C - D_H) - \beta_0(H)]] \end{aligned}$$

such that the following incentive constraints, participation constraints and truth telling constraints are satisfied.

<u>Incentive Constraints</u>

$$\begin{aligned} &\beta_D(j) - \frac{\beta_Y(j)\theta_j\, e_j}{C - D_j} && = 0 \\ &\beta_Y(j)\,\theta_j \ln(C - D_j) && = M\, e_j \quad j = L, H \end{aligned} \tag{IC´}$$

$$\beta_D\left(\hat{j}\right)-\frac{\beta_Y\left(\hat{j}\right)\theta_j\, e_{\hat{j}j}}{C-D_{\hat{j}j}} \qquad = 0$$

$$\beta_Y\left(\hat{j}\right)\theta_j \ln(C-D_{\hat{j}j}) \qquad = M\, e_{\hat{j}j} \quad j=L,H;\hat{j}=L,H;j\neq\hat{j} \tag{IC´´}$$

<u>Participation Constraints</u>

$$\beta_0(j)+\beta_D(j)D_j+\beta_Y(j)\theta_j\, e_j \ln\left(C-D_j\right)-\frac{1}{2}M\, e_j^2 \geq \omega_0 \quad j=L,H \tag{PC´}$$

where ω_0 is the reservation wage.

<u>Truth telling Constraints</u>

$$\begin{aligned} &E\left[\beta_0(j)+\beta_D(j)D_j+\beta_Y(j)\tilde{Y}-\frac{1}{2}Me_j^2\,\middle|\,\theta_j,D_j,e_j\right] \\ &\geq E\left[\beta_0\left(\hat{j}\right)+\beta_D\left(\hat{j}\right)D_{\hat{j}j}+\beta_Y\left(\hat{j}\right)\tilde{Y}-\frac{1}{2}Me_{\hat{j}j}^2\,\middle|\,\theta_j,D_{\hat{j}j},e_{\hat{j}j}\right] \\ &\qquad\qquad j=L,H;\hat{j}=L,H;j\neq\hat{j} \end{aligned} \tag{TT}$$

The incentive constraint (IC´) in Program 1A is obtained by substituting the first order conditions from the agent's optimization problem in the incentive constraint (IC) of Program 1. The participation constraint (PC´) is obtained by explicitly writing out the conditional expected compensation. These constraints are derived for the case where the agent reports his/her type truthfully. The incentive constraint (IC´´) is derived in a similar manner and is applicable to those cases in which the agent falsifies his/her type. The truth inducing constraints (TT) makes use of the revelation principle and ensures that in equilibrium the agent reports his/her type truthfully. Note that we need to include (IC´´) in the incentive constraints to take into account the off-equilibrium behaviour. (IC´´) describes the choice of agent when the agent send a false report of his/her type.

In equilibrium however the agent declares his/her type truthfully and (IC´´) is not binding.

The solution to Program 1A will determine the equilibrium values for the fixed component of compensation offered to each agent [$\beta_0(j)$, j=H,L], the component of compensation linked to the dividend payment [$\beta_D(j)$, j=H,L], the component of compensation linked to the final output [$\beta_Y(j)$, j=H,L], the amount of dividend declared by each agent [D_j, j=H,L], and the quantum of effort supplied by each agent [e_j, j=H,L]. The principal will choose [$\beta_0(j)$, $\beta_D(j)$, $\beta_Y(j)$]; the agent will choose $D(j|\beta_D(j), \beta_Y(j))$ and $e(j|\beta_D(j), \beta_Y(j))$.

Our interest is in the information that is transmitted in the process of solving the principal/agent problem and hence we focus on a separating equilibrium. For a separating contract to be an equilibrium it must be the case that the payoff to the principal is larger from a separating contract than it would be from a pooling contract. This in turn will place restrictions on the model parameters, especially the probability distribution of types. We assume these restrictions hold so that our analysis deals only with separating contracts. We propose to solve this problem by the following method.

Step 1: Find the agent's optimal choice of D and e for a particular value of the coefficients of the wage function. These can be found by solving the equations in the incentive condition.

Step 2: Substitute these optimum values of D and e in the agent's objective function and the principal's objective function. This will express the

agent's utility and the principal's utility in terms of the coefficient values.

Step 3: We find the indifference curves for the principal and agent from the expression obtained in Step 2. The indifference curves are drawn in the space (β_Y, β_D).

Step 4: Identify a contract for the lowest type L on the indifference curve corresponding to the binding participating wage ω_0.

Step 5: Given the contract for L, we now identify the space in which the separating contracts for H must lie. We do not comment at this stage on the nature of the equilibrium contract. We just use the *single crossing criterion* in order to delineate the contracts in (β_Y, β_D) space.[4]

Step 6: Maximize Principal's expected profit in the space identified in Step 5.

The approach can be illustrated by referring to Figure 1 and Figure 2. We show in the appendix that the indifference curves in the (β_Y, β_D) plane for both the H-type and L-type agents are concave and downward sloping and that the sorting condition is met. We can further elaborate these ideas by referring to Figure 1. In this figure we have drawn the indifference surfaces for the L-type and the H-type agent. The indifference surfaces are drawn in the space (β_0, β_D, β_Y). Figure 2 shows the cross-sectional view of the indifference surfaces drawn in Figure 1, the cross-sectional plane running parallel to the (β_Y, β_D) plane.

---

[4]The single crossing criterion is also known as the Spence-Mirrless criterion or the sorting criterion. It implies that the indifference curves for two type of agents can cross only once. For details see Tirole (1988) and Fudenberg and Tirole (1991) .

In Figure 1, $B_L LL'$ (thin line) is the indifference surface for L for a net expected reservation wage of ω_0 (net of the cost of effort). Suppose W_L is the contract offered to L. $B_H HH'$ (thick line) is the indifference surface for H passing through W_L . OE (the dashed line) is the line on the (β_D, β_Y) plane, such that $\beta_D = \beta_Y$.

In Figure 2, LL' is the section of the indifference surface for L and HH' is the section of the indifference surface for H. The separating contract offered to H must lie in the area bounded by section $W_L H'$ and $W_L L'$ of the indifference plane (see Figure 2). Consider a contract represented by point A. If this contract is offered then both L-type and the H-type would prefer the contract W_L to A. Conversely, if the contract B is offered, then both the L-type and H-type would prefer the contract B to the contract W_L. In these cases, the contract will not be separating.

Let ω_H be the utility that H would get by accepting the contract offered to L [in Figure 1 it would be the utility obtained on the indifference surface for H] . Clearly H would reject any contract that offers him/her an utility less than ω_H. It can be shown that $\omega_H > \omega_0$. The L-type agent is paid the reservation wage, while the H-type agent is paid an information rent equal to ($\omega_H - \omega_0$).

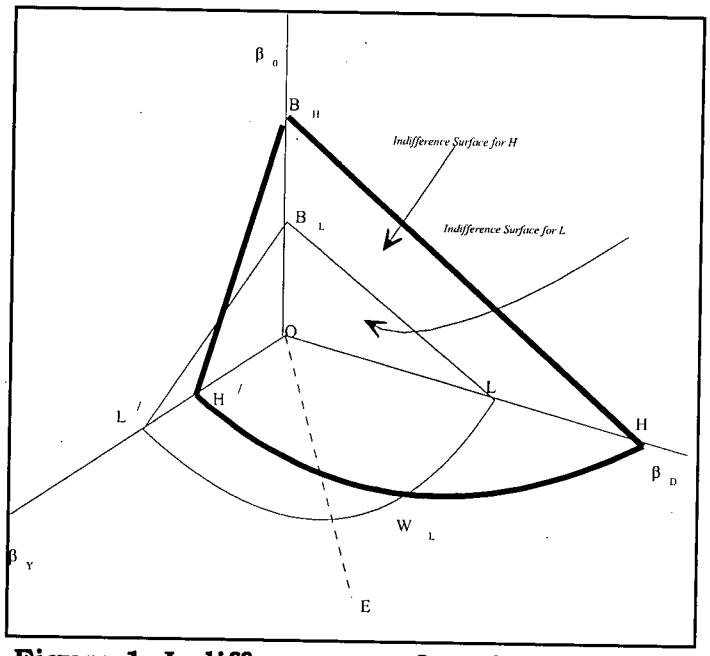

Figure 1: Indifference surface for H and for L in the (β_0, β_D, β_Y) space.

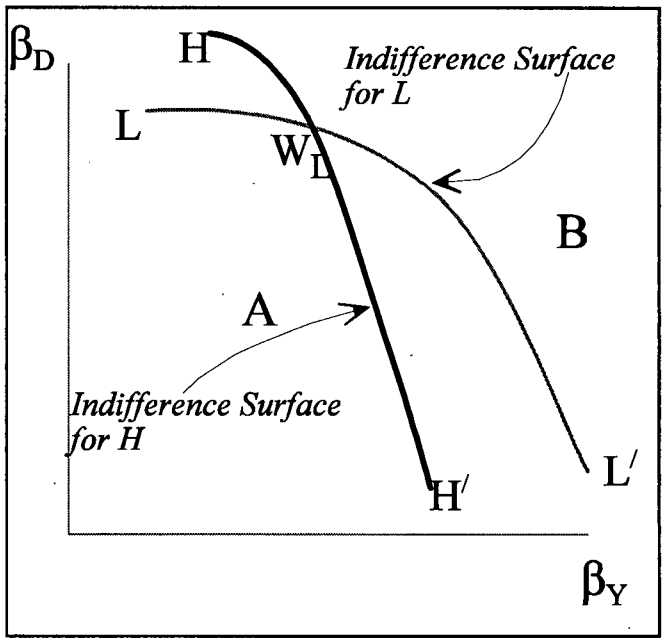

Figure 2: Cross Sectional View of the Indifference Surfaces for H and L — Cross section by a plane parallel to β_D and β_Y .

Since the H-type is restricted on the indifference surface, having a value ω_H, the principal should rationally extract the first-best output. This can be done by setting $\beta_D(H) = \beta_Y(H) = 1$ and setting $\beta_0(H)$ in such a way so as to make the ex-ante utility to H equal to ω_H. The contract offered to L would be an inefficient contract which would optimize the expected gain to the principal. However, the more efficient the contract offered to L is, the higher is the rent paid to H. The principal therefore balances the efficiency lost by offering an inefficient contract to L against the rent paid to H. The contract offered to H should lie on the indifference surface for H (viz., on the surface B_HHH') and, in addition, it must lie between two planes: one is the plane containing the β_0 axis and passing through points O and W_L; the other is the plane defined by the β_Y and β_0 axis. If the contract offered to H is in this space, then L would strictly prefer his/her own contract W_L over the contract offered to H and H would be indifferent between the contract offered to him/her and the contract offered to L.

By looking at the structure of the equations (1) and (2), we know that H can be induced to achieve the first-best result by setting $\beta_D(H) = \beta_Y(H) = 1$. The principal will design the contract W_L so as to maximize the expected gain to himself/herself. In general, we cannot expect the contract W_L to be such as to induce efficient (first best) decisions for L, because in order to induce efficient decisions, L will also have to be offered a contract with $\beta_D(L) = \beta_Y(L) = 1$. In that case the contract W_L will be on the line OE. However, this means that the rent paid to H (which is determined by the indifference surface for H passing through W_L) will be higher. Therefore, the principal will maintain balance between the

rent paid to H and the inefficient decision by L, so as to maximize his/her expected profit.[5] We can also see that the contract W_L will be such that the line joining O and W_L must remain in the region OE and the β_D axis. This is because otherwise L would prefer the efficient contract given to H and in the process derive a rent.

This implies that $\frac{\beta_D(L)}{\beta_Y(L)} > 1$. We also note that $\beta_Y(L)<1$, because otherwise H would prefer the contract W_L as opposed to the efficient contract.

We summarize the conclusions reached so far with the following proposition.

Proposition 1

The L-type will be offered an inefficient contract such that the ex-ante net expected payment to L-type is equal to the reservation wage ω_0. The contract offered to L would be such that $\frac{\beta_D(L)}{\beta_Y(L)} > 1$ and $\beta_Y(L) < 1$ and $\beta_0(L)$ is such that the ex-ante net gain of agent L is equal to the reservation wage ω_0.

The H-type would be offered an efficient contract to induce first-best results. This would be achieved by setting $\beta_D(H) = \beta_Y(H) = 1$. $\beta_0(H)$ would be set at a level so that the ex-ante net expected payment to H is the same, irrespective of whether he accepts this contract or the contract offered to L.

We can now establish the main result of our dissertation.

---

[5]The result that the more productive type is offered an efficient contract (i.e one which achieves first best outcomes) and the less productive type is offered an inefficient contract so as to reduce the information rent paid to the more productive agent, is a standard result in the principal-agent literature. For example, see Laffont and Tirole (1993) and Varian (1992).

Proposition 2

Assuming interior solutions to the Agent's problem, the H-type agent will declare a smaller dividend than the L-type agent. The H-type agent will also put in more effort than the L-type agent.

Proof: See Appendix I

We can use Figure 3 to illustrate that the dividend declared by L would be higher than the dividend declared by H.

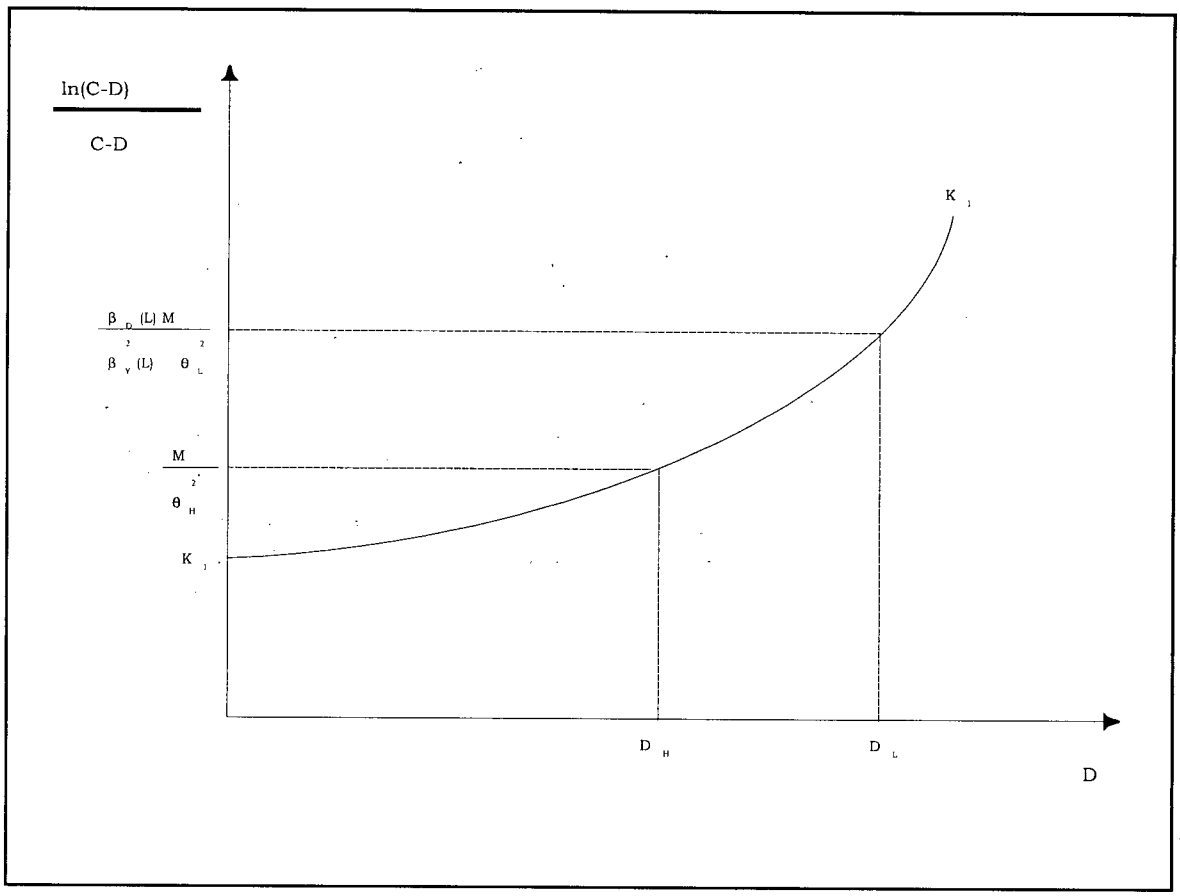

Figure 3: Dividend decision by the agent

Recall that the dividend selected by a particular type is governed by the first order conditions of the agent's optimization problem. These first order conditions are:

$$\beta_D(j) - \frac{\beta_Y(j)\theta_j e_j}{C - D_j} = 0$$

$$\beta_Y(j)\theta_j \ln(C - D_j) = M e_j$$

Eliminating e_j, we get

$$\frac{\ln(C - D_j)}{(C - D_j)} = \frac{\beta_D(j)}{\beta_Y^2(j)}\frac{M}{\theta_j^2} \qquad j = L, H$$

$$= \frac{M}{\theta_H^2} \qquad for\ H$$

$$= \frac{\beta_D(L)}{\beta_Y^2(L)}\frac{M}{\theta_L^2} \qquad for\ L$$

In Figure 3, D is measured along x-axis. The curve K_1K_1 plots $\frac{\ln(C_1 - D)}{C_1 - D}$. The dividend declared by H will be that value of D for which $\frac{\ln(C_1 - D)}{C_1 - D} = \frac{M}{\theta_H^2}$. So, then, we measure a distance equal to $\frac{M}{\theta_H^2}$ along the y-axis and draw a horizontal line through it. The point at which this line intersects the curve K_1K_1 will give us the dividend declared by H.[6]

Similarly, we measure $\frac{\beta_D(L)}{\beta_Y^2(L)}\frac{M}{\theta_L^2}$ along the vertical axis and draw a horizontal line at that distance. The point of intersection of this line with curve

[6]It should be noted that the curve K_1K_1 would be an increasing function of (C–D) only if C–D>Exp(1) where Exp(1) is the exponential function. As stated in footnote3, this condition will have to be satisfied for an interior optimal solution.

K_1K_1 will give us the dividend declared by L. The dividend declared by L will be more than the dividend declared by H because $\frac{\beta_D(L)}{\beta_Y^2(L)}\frac{M}{\theta_L^2} > \frac{M}{\theta_H^2}$.

The result is intuitive. The earnings from the previous period are available for apportionment into investment and dividend. The shareholders want the agent with the higher productivity to invest more than the agent with the lower productivity. They achieve such an end by offering two separating contracts. The agent with lower productivity picks the contract with the higher weight on dividend and declares a higher dividend. The agent with higher productivity, on the other hand, picks the efficient contract and invests more in the production process. As a result, the agent with higher productivity declares a smaller dividend than the agent with lower productivity.

We can use Figure 4 to understand the comparative statics of changing the initial availability of cash. Curve K_2K_2 shows the way curve K_1K_1 shifts if the cash availability is increased from C_1 to C_2. We notice that the curve shifts to the right, and that the dividend level increases for both H and L. This helps us understand the impact of a change in earnings upon dividends. With an increase in earnings, therefore, dividends will increase for both the type of agents, although Type H will still declare a lower dividend than Type L.

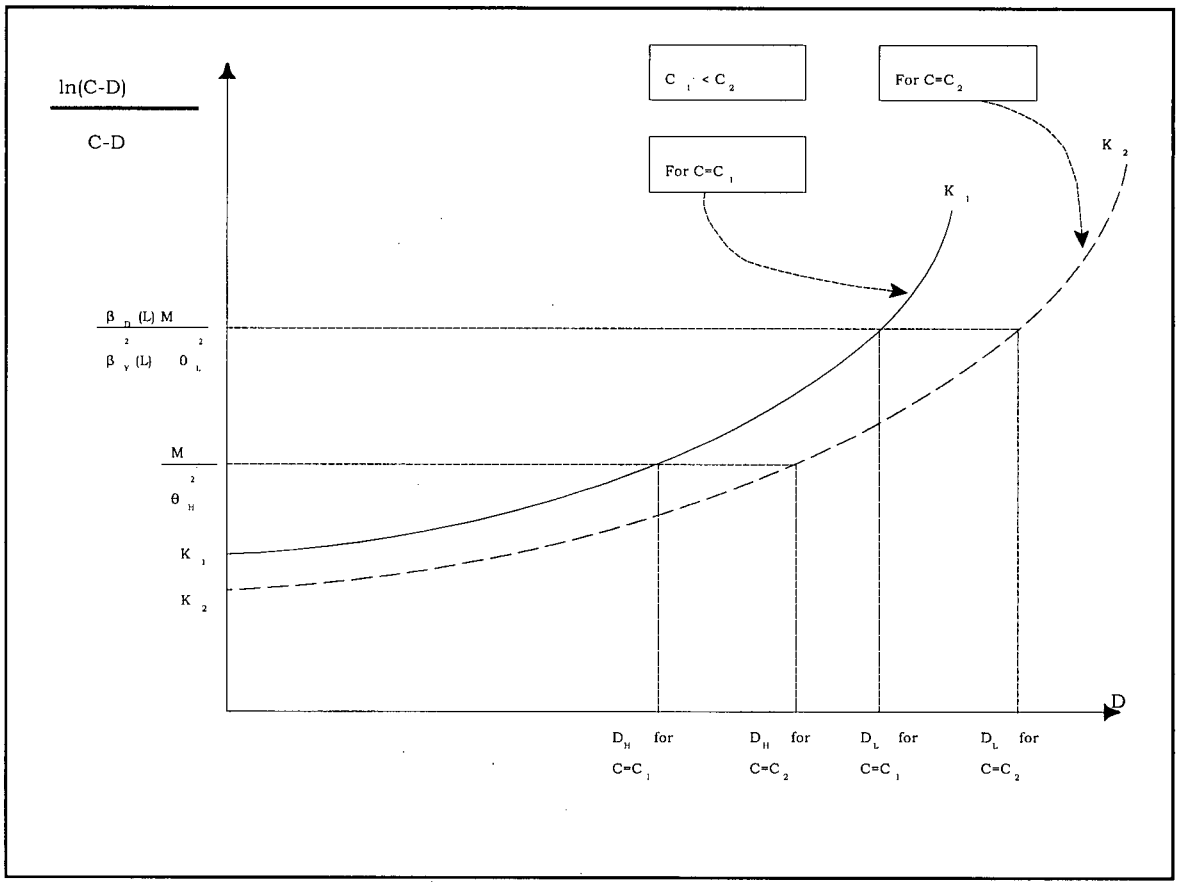

Figure 4: Impact of initial cash availability on dividends paid.

3.7 When Difference In Cash Availability Characterize The Difference In Agents

So far we have characterized the difference in the agents by the productivity parameter θ_j, j = H,L, in a model in which C, the cash available, is the same for both the agents. That is, we have assumed that the asymmetric information between the agent and the principal is about the productivity parameter θ_j, j = H,L, and that there is no asymmetric information about C, the cash available. We now switch the source of information asymmetry, characterize the difference

between agents by the difference in cash availability, and bestow the agents with the same productivity parameter θ. Specifically, we assume that both the H-type and L-type agents have the same productivity parameter θ, and that this parameter value is common knowledge. We further assume that the cash available, C_j, j = H,L, is different for the two agents and $C_H > C_L$. We also assume that the true value of C is known only to the agent. The problem statements for this case are similar to the problem statements we have considered up until now, except that θ_j is replaced with θ and C would be replaced with C_j. We notice that in this case although the single crossing property remains valid, at any point on the (β_D, β_Y) plane the slope of the indifference curve for H is lower than the slope of the indifference curve for L.

In Figure 5 we have drawn the indifference surfaces for the L-type and the H-type agent. The indifference surfaces are drawn in the space (β_0, β_D, β_Y). In Figure 5, B_LLL' (thin line) is the indifference surface for L, given a net expected reservation wage of ω_0 (net of the cost of effort). Suppose W_L is the contract offered to L. B_HHH' (thick line) is the indifference surface for H passing through W_L. OE (the dashed line) is the line on the (β_D, β_Y) plane, such that $\beta_D = \beta_Y$.

Figure 6 shows the cross-sectional view of the indifference surfaces drawn in Figure 5, with the cross-sectional plane aligned parallel to the (β_D, β_Y) plane. In Figure 6, LL' is the section of the indifference surface for L, and HH' is the section of the indifference surface for H. The separating contract offered to H must lie in the area bounded by the section of the indifference plane W_L H and W_LL (see Figure 6). Consider a contract represented by point A: if this contract is

offered, then both the L-type and the H-type would prefer contract W_L to contract A. Conversely, if the contract B is offered, then both the L-type and the H-type would prefer contract B to contract C_L. Under these circumstances, the contract will not be separating. We can see that the region for separating contract is different in these two cases, and in a sense, this region flips from being below W_L to being above W_L. A straightforward extension of the argument used to enunciate Proposition 1 can now be used to enunciate the following corresponding proposition.

Proposition 3

This proposition is for the case where agents are characterised by the difference in their initial earning E. The L-type is offered an inefficient contract such that the ex-ante net expected payment to L-type is equal to the reservation wage ω_0. The contract offered to L is such that $\frac{\beta_D(L)}{\beta_Y(L)} < 1$ *and* $\beta_Y(L) < 1$ *and* $\beta_0(L)$ *is such that the ex-ante net gain of agent L is equal to the reservation wage* ω_0.

The H-type is offered an efficient contract to induce first best results. This is achieved by setting $\beta_D(H) = \beta_Y(H) = 1$. $\beta_0(H)$ *is set at a level so that the ex-ante net expected payment to H is irrespective of whether he accepts this contract or the contract offered to L*

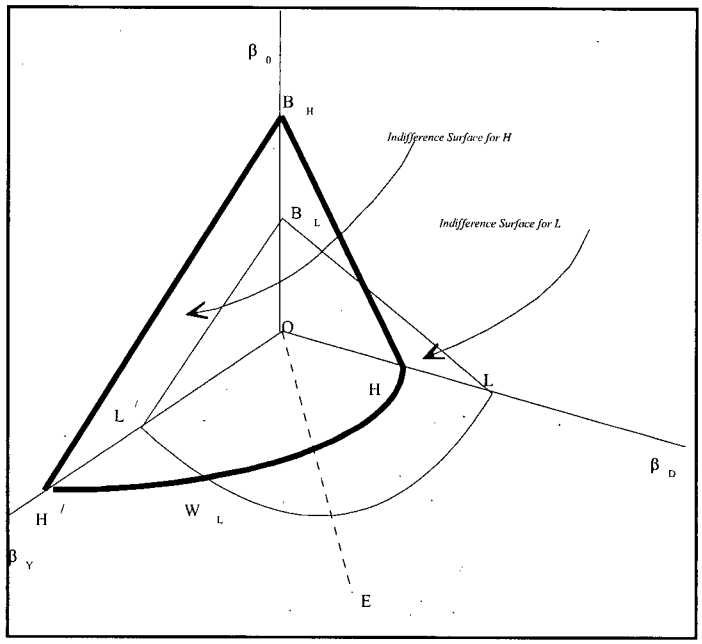

Figure 5 : Indifference surface for H and for L in the (β_0, β_D, β_Y) space when agents are differentiated by cash availability .

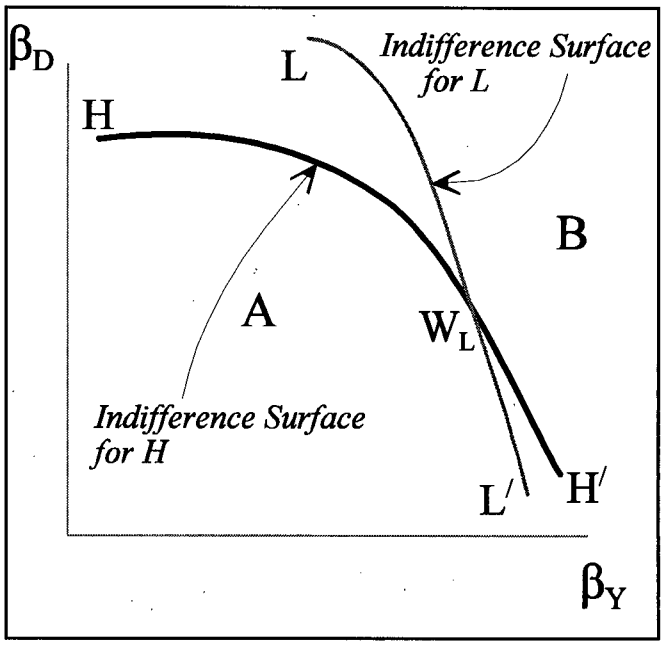

Figure 6: Cross Sectional View of the Indifference Surfaces for H and L when agents are differentiated by cash availability — Cross section by a plane parallel to β_D and β_Y .

By comparing **Proposition 1** with **Proposition 3, ?** we can see that in both scenarios, H-type is offered an efficient contract while L-type is offered an inefficient contract. However, the nature of the inefficient contract offered to L is different in each case. Where the difference between the agents is characterized by productivity parameter θ (i.e., different agents have different θ, but same prior earning E), the inefficient contract offered to L is such that $\frac{\beta_D(L)}{\beta_Y(L)} > 1$. On the other hand, when the difference between the agents is characterized by cash availability C (i.e different agents have different C but the same productivity parameter), the inefficient contract offered to L is such that $\frac{\beta_D(L)}{\beta_Y(L)} < 1$. Given this difference in the contract offered to L, we know intuitively that the dividend behaviour would be different in the two cases. We can show that, in the latter case, H will give a higher dividend than L if the investment level is high.

Proposition 4

This proposition is for the case in which agents are characterized by the difference in their initial cash C. Assuming interior solutions to the agent's problem, if the investment level is high then the H-type agent will declare a larger dividend than the L-type agent.

Proof: See Appendix I

3.8 Discussion

We have seen two cases:

Case 1: When difference between agents is characterized by the difference in productivity; and

Case 2: When difference between agents is characterized by the difference in the initial cash availability.

We find that when agents differ in their productivity, then, *for a given level of cash*, the agent with higher productivity pays a smaller dividend. In contrast, when agents have the same productivity, yet differ in their initial cash availability, then the agent with the higher initial cash availability pays the higher dividend.

These results are intuitive. When the agents differ in their productivity, then the agent with the higher productivity can generate more expected return per dollar of investment. Therefore, *for a given level of cash*, the agent with higher productivity will invest more, and, consequently, will have less money left over to pay as a dividend. On the other hand, an agent with less productivity would invest less because the expected return he/she can generate per dollar of investment is less than the corresponding return generated by the agent with higher productivity. The agent with less productivity would therefore have more cash to pay out as a dividend.

In contrast, when the agents have the same productivity, and differ only in their initial cash availability, then both types of agent have the same expected rate of return per dollar of investment. As a result, in the first-best world, both the agents invest the same amount and exert the same effort. The agent with the higher type pays a greater dividend because he/she has a greater cash endowment, and has more money left over to pay the dividend. In the second best world, too, the agent with the higher cash endowment pays a higher dividend and invests optimally. The agent with the smaller cash endowment overinvests and pays less dividend.

Chapter 4

Generalized Production Function and Continuum of Agent Types

This chapter's setting closely resembles the previous chapter's, except that here we use a generalized production function and a generalized cost of effort function. This model is based on an extension of the model in Paik and Sen (1995)[7]. They focus on issues of capital rationing in a divisionalized setting.

We assume that the agent type θ is distributed continuously over a support [θ_{Min}, θ_{Max}]. By definition, θ is higher for an agent with a higher level of skill. The agent knows his own θ; however, the principal knows only the distribution of θ. The principal has a prior belief about the probability density of θ, which is represented by the probability density function f(θ).

The stochastic output (production) function is given by

$$\tilde{Y} = X(e, I, \theta) + \tilde{\varepsilon} \text{ where } \tilde{\varepsilon} \text{ is random noise} \tag{3}$$

As in the previous section, e is the effort supplied by the agent and I is the investment and I=C-D where D is the dividend. We posit that the choice variables e and I are determined by an optimal contract and that e and I are functions of

[7]My model differs from the model of Paik and Sen in the following aspects: a)They don't have dividends in their model; b) they don't have convex, cost of effort function in their model; c) they don't sign the direction of change in dividend and effort with agent type; d) in their model principal supplies the investment after agent type is declared whereas in our model cash is in place before the contract is signed and the agent determines the investment level.

agent type θ . $X(\bullet)$ is the expected output from the agent's point of view. The conditional distribution of $\tilde{Y}$, conditioned on $X(e,I,\theta)$ is represented by $g(\tilde{Y} \mid X(e,I,\theta))$ or, in short, by $g(\tilde{Y} \mid X(\bullet))$. We assume that $X(\bullet)$ is such that $X_e(\bullet)>0$ and $X_\theta(\bullet)>0$. We also assume constant support for $\tilde{Y}$.

As in the previous section, both principal and agent are risk-neutral. The agent has a disutility for effort or a private cost for effort. This cost is represented as $k(e)$. We assume that $k'(e)>0$, $k''(e)>0$; that is that the cost function is increasing and convex in effort.

The principal offers a menu of wage contracts to the agent. We denote a wage contract by $\omega(\tilde{Y}, D, \hat{\theta})$. The wage contract is a function of three variables: two observable variables, namely output (production) at the end of period, and dividend declared at the beginning of the period; and one message about the type of agent.

$U[e, D, \hat{\theta}, \theta]$ is the agent's expected utility if the agent is of type θ and declares himself/herself to be of type $\hat{\theta}$.

We see that

$$U\left[e, D, \hat{\theta}, \theta\right]$$

$$= \int_{-\infty}^{+\infty} \omega\left[\tilde{Y}, D, \hat{\theta}\right] g\left[\tilde{Y} \middle| X\left[e, I, \theta\right]\right] dy - k(e)$$

The principal's problem can now be represented as in Program 2 below.

Program 2

$$\underset{\omega(\bullet),e(\theta),D(\theta)}{Maximize} \int_{\theta_{Min}}^{\theta_{Max}} \Big[D(\theta)+X\big[e(\theta),I(\theta),\theta\big] - \int_{-\infty}^{+\infty} \omega\big[\tilde{Y},D(\theta),\theta\big]g\big[\tilde{Y}\big|X\big[e(\theta),I(\theta),\theta\big]\big]dy\Big]f(\theta)d\theta$$

such that

$$\begin{bmatrix}\theta\\ e(\theta)\\ D(\theta)\end{bmatrix} \in \underset{\hat{\theta},e,D}{Arg\max}\left[\int_{-\infty}^{+\infty}\omega\big[\tilde{Y},D,\hat{\theta}\big]g\big[\tilde{Y}\big|X\big[e,I,\theta\big]\big]dy - k(e)\right] \qquad \forall\theta \qquad \text{(ICG)}$$

and

$$U\big(e(\theta),D(\theta),\theta,\theta\big) \geq \omega_0 \qquad \forall\theta \qquad \text{(PCG)}$$

where $D(\theta) = C - I(\theta) \geq 0 \ \forall\theta$. This equation implies that in this model we refrain from allowing the agent to raise money from the capital market through debt or equity issue. The dividend, in our model, is the excess of earnings over investment. We assume that the non-negativity condition for D(θ) is satisfied .

The first constraint (ICG) is the incentive constraint; the second (PCG) is the participation constraint. Following the analysis in Paik and Sen (1995), we substitute the agent's global incentive compatibility condition with a local incentive compatibility condition that ensures that the agent tells the truth (i.e., declares his type correctly) and spends the optimal amount of effort.

We assume that the optimal solution is an interior one and that the first order approach to generalized principal-agent problem is valid. The first order condition for optimal effort is

$$\begin{aligned} &\frac{\partial}{\partial e} U\left[e\left(\hat{\theta},\theta\right), D\left(\hat{\theta},\theta\right), \hat{\theta}, \theta\right] = 0 \\ \Rightarrow\ & X_e[\bullet] \int_{-\infty}^{+\infty} \omega\left[\tilde{Y}, D, \hat{\theta}\right] g_X\left[\tilde{Y} \middle| X[e, I, \theta]\right] dy = k'(e) \forall \theta, \hat{\theta} \end{aligned} \tag{4}$$

The local condition for agent's truth-telling is

$$\frac{\partial}{\partial \hat{\theta}} U\left[e\left(\hat{\theta},\theta\right), D\left(\hat{\theta},\theta\right), \hat{\theta}, \theta\right] = 0 \tag{5}$$

In order to elicit a truthful report from the agent, the principal will have to pay the agent enough so that the agent's equilibrium utility increases in θ.

To see this, we differentiate totally the agent's equilibrium utility U[θ] = U[e(θ),D(θ),θ,θ] with respect to the true type θ.

$$\begin{aligned} \frac{d}{d\theta} U\left[e(\theta), D(\theta), \theta, \theta\right] \quad = \quad & \frac{\partial}{\partial e} U\left[e\left(\hat{\theta},\theta\right), D\left(\hat{\theta},\theta\right), \hat{\theta}, \theta\right]\Big|_{\hat{\theta}=\theta} \frac{de}{d\theta} \\ + \ & \frac{\partial}{\partial D} U\left[e\left(\hat{\theta},\theta\right), D\left(\hat{\theta},\theta\right), \hat{\theta}, \theta\right]\Big|_{\hat{\theta}=\theta} \frac{dD}{d\theta} \\ + \ & \frac{\partial}{\partial \hat{\theta}} U\left[e\left(\hat{\theta},\theta\right), D\left(\hat{\theta},\theta\right), \hat{\theta}, \theta\right]\Big|_{\hat{\theta}=\theta} \frac{d\hat{\theta}}{d\theta} \\ + \ & \frac{\partial}{\partial \theta} U\left[e\left(\hat{\theta},\theta\right), D\left(\hat{\theta},\theta\right), \hat{\theta}, \theta\right]\Big|_{\hat{\theta}=\theta} \end{aligned}$$

The first three terms on the right hand side vanish because of first order conditions for an interior optima and because of the truth telling condition.

$$\begin{aligned}\therefore \frac{d}{d\theta}U\left[e(\theta),D(\theta),\theta,\theta\right] &= \frac{\partial}{\partial\theta}U\left[\bullet\right] \\ &= X_\theta\left[\bullet\right]\int_{-\infty}^{+\infty}\omega\left[\tilde{Y},D(\theta),\theta\right]g_X\left[\tilde{Y}\middle|X\left[e(\theta),I(\theta),\theta\right]\right]dy \\ &= \frac{X_\theta\left[\bullet\right]}{X_e\left[\bullet\right]}k'\left(e(\theta)\right) \quad \text{from optimal effort condition.} \\ &>0\end{aligned}$$

Since U(θ) increases with θ, the participation constraint may be replaced by a single constraint for the worst type θ_{Min}. The participation constraint therefore becomes $U(\theta_{Min}) = \omega_0$.

We write $\dfrac{X_\theta\left[\bullet\right]}{X_e\left[\bullet\right]} \equiv \gamma(\theta)$.

$$\begin{aligned}\therefore \frac{d}{d\theta}U\left[e(\theta),D(\theta),\theta,\theta\right] &= \frac{X_\theta\left[\bullet\right]}{X_e\left[\bullet\right]}k'\left(e(\theta)\right) \\ &= \gamma(\theta)k'\left(e(\theta)\right)\end{aligned}$$

$$\begin{aligned}\Rightarrow U\left[e(\theta),D(\theta),\theta,\theta\right] &= \int_{\theta_{Min}}^{\theta}\gamma(s)k'\left(e(s)\right)ds+\omega_0 \\ &= \int_{-\infty}^{+\infty}\omega\left[\tilde{Y},D(\theta),\theta\right]g\left[\tilde{Y}\middle|X\left[e(\theta),I(\theta),\theta\right]\right]dy-k\left(e(\theta)\right)\end{aligned}$$

from the definition of $U[\theta]$

(6)

We can now rewrite Program 2 as Program 2A by incorporating (6).

Program 2A

$$\underset{\omega(\bullet),e(\theta),D(\theta)}{Maximize} \int_{\theta_{Min}}^{\theta_{Max}} \Big[D(\theta)+X\big[e(\theta),I(\theta),\theta\big]$$

$$-\int_{-\infty}^{+\infty} \omega\big[\tilde{Y},D(\theta),\theta\big]g\Big[\tilde{Y}\Big|X\big[e(\theta),I(\theta),\theta\big]\Big]dy\Big]f(\theta)d\theta$$

such that

$$\int_{-\infty}^{+\infty} \omega\big[\tilde{Y},D(\theta)\big]g\Big[\tilde{Y}\Big|X\big[e(\theta),I(\theta),\theta\big]\Big]dy-k\big(e(\theta)\big)$$

$$=\int_{\theta_{Min}}^{\theta} \gamma(s)k'\big(e(s)\big)ds+\omega_0 \qquad \forall\ \theta$$

We simplify the objective function in Program 2A by incorporating the constraint into the objective function and simplifying. Note that the incorporation of the constraint into the maximand eliminates the wage function ω(•) from the maximand. We now maximize with respect to D(•) and e(•), because the principal now induces these choices through the choice of an optimal contract. The objective function now becomes

$$\underset{e(\theta),D(\theta)}{Maximize} \int_{\theta_{Min}}^{\theta_{Max}} \Big[D(\theta)+X\big[e(\theta),I(\theta),\theta\big]-k\big(e(\theta)\big)\Big]f(\theta)d\theta$$

$$-\int_{\theta_{Min}}^{\theta_{Max}} \Big[\int_{\theta_{Min}}^{\theta} \gamma(s)k'\big(e(s)\big)ds\Big]f(\theta)d\theta-\omega_0$$

$$\equiv \underset{e(\theta),D(\theta)}{Maximize} \int_{\theta_{Min}}^{\theta_{Max}} \Big[D(\theta)+X\big[e(\theta),I(\theta),\theta\big]-k\big(e(\theta)\big)\Big]f(\theta)d\theta$$

$$-\int_{\theta_{Min}}^{\theta_{Max}} \Big[\int_{\theta}^{\theta_{Max}} \big[f(\tau)d\tau\big]\gamma(s)k'\big(e(s)\big)ds\Big]-\omega_0$$

by changing the order of integration in the second integral

$$\equiv \quad \underset{e(\theta),D(\theta)}{Maximize} \int_{\theta_{Min}}^{\theta_{Max}} \Big[D(\theta)+X\big[e(\theta),I(\theta),\theta\big]-k\big(e(\theta)\big)$$

$$-\gamma(\theta)k'\big(e(\theta)\big)h(\theta)\Big]f(\theta)d\theta-\omega_0$$

where $h(\theta)=\dfrac{1-F(\theta)}{f(\theta)}$ is the inverse hazard function

Pointwise optimization characterizes optimal choice of e and I (also we reiterate here that D=C-I). Thus we have

$$\begin{aligned}
& -1+X_I[\bullet]-\gamma_I(\bullet)k'(\bullet)h(\theta)=0 \\
& X_e[\bullet]-k'(\bullet)-\big[\gamma_e(\bullet)k'(\bullet)+\gamma(\bullet)k''(\bullet)\big]h(\theta)=0 \\
\Rightarrow \quad & X_I[\bullet]=1+\gamma_I(\bullet)k'(\bullet)h(\theta) \\
& X_e[\bullet]=k'(\bullet)\big[1+\gamma_e(\bullet)h(\theta)\big]+\gamma(\bullet)k''(\bullet)h(\theta)
\end{aligned} \tag{7}$$

The optimal e(θ) and I(θ) [and hence D(θ)] are jointly determined by the two partial differential equations in (7). The optimal contract is characterized by (6).

Proposition 5

In the optimal second-best solution, both effort and investment are less than in the first-best case except at $\theta = \theta_{Max}$.

Proof: The first-best solution will be given by $X_I(\bullet) = 1$ and $X_e(\bullet) = k'(e)$. Comparing the first-best with the second-best, we notice that, in the second-best case [as given in (7)], the optimal marginal product for effort and for investment is more. Therefore, the second-best solution provides less marginal effort and investment than the first-best.■

In order to understand the change in investment I (and, equivalently, the change in dividend D) and effort e as the productivity parameter θ changes, we

need to impose some structure on the production function X(•), the cost of effort function k(e), and the probability distribution followed by θ. We consider the following two cases:

Case 1: $X(\bullet)$ satisfies the following conditions.

$$X_I(\bullet) > 0,\ X_{ee}(\bullet) < 0,\ X_{II}(\bullet) < 0,\ X_{ee}(\bullet)X_{II}(\bullet) \geq X_{eI}^2(\bullet),\ X_{eI}(\bullet) \geq 0$$

$$X_{\theta e}(\bullet) > 0, X_{\theta I}(\bullet) > 0, \gamma_I(\bullet) = \gamma_\theta(\bullet) = 0,\ \gamma_e(\bullet) = \text{Constant (Say } \kappa) \geq 0$$

k(e) satisfies the condition $k'''(e) \geq 0$. h(θ) satisfies the condition $h'(\theta) \leq 0$.[8]

Applying these assumptions to (7), we get

$$\begin{aligned} X_I[\bullet] &= 1 \\ X_e[\bullet] &= k'(\bullet)\left[1 + \kappa h(\theta)\right] + \gamma[\bullet]k''(\bullet)h(\theta) \end{aligned} \tag{8}$$

Totally differentiating (8) with respect to θ , we get

$$X_{II}[\bullet]\frac{dI}{d\theta} + X_{Ie}[\bullet]\frac{de}{d\theta} + X_{I\theta}[\bullet] = 0$$

$$X_{Ie}[\bullet]\frac{dI}{d\theta} + \left[X_{ee}[\bullet] - k''(\bullet)\left\{1 + 2\kappa h(\theta)\right\} - \gamma[\bullet]k'''(\bullet)h(\theta)\right]\frac{de}{d\theta}$$
$$+\left[X_{e\theta}[\bullet] - h'(\theta)\left\{\kappa k'(\bullet) + \gamma[\bullet]k''(\bullet)\right\}\right] = 0$$

Solving, we get

---

[8]These assumptions about the production function X(•) imply that X(•) is weakly concave with respect to e and I. An example of the production function which satisfies these conditions is $X(e, I, \theta) = Exp\ (\theta)e^{\tau} I^{(1-\tau)}$, $0 < \tau < 1$. This is a Cobb-Douglas production function. Quadratic cost functions satisfy the structure specified for cost functions, and exponential distributions satisfy the condition on distribution of θ.

$$\frac{\dfrac{dI}{d\theta}}{\left[X_{Ie}[\bullet]\left\{X_{e\theta}[\bullet]-h'(\theta)\left\{\kappa k'(\bullet)+\gamma[\bullet]k''(\bullet)\right\}\right\}\right.}$$

$$\left.-X_{I\theta}[\bullet]\left\{X_{ee}[\bullet]-k''(\bullet)\left\{1+2\kappa h(\theta)\right\}-\gamma[\bullet]k'''(\bullet)h(\theta)\right\}\right]$$

$$=\frac{\dfrac{de}{d\theta}}{X_{I\theta}[\bullet]X_{Ie}[\bullet]-X_{II}[\bullet]\left\{X_{e\theta}[\bullet]-h'(\theta)\left\{\kappa k'(\bullet)+\gamma[\bullet]k''(\bullet)\right\}\right\}}$$

$$=\frac{1}{X_{II}[\bullet]\left\{X_{ee}[\bullet]-k''(\bullet)\left\{1+2\kappa h(\theta)\right\}-\gamma[\bullet]k'''(\bullet)h(\theta)\right\}-X_{Ie}^{2}[\bullet]}$$

From the calculations above we can determine that under the scenario represented by our assumptions, $\frac{dI}{d\theta}>0$. This implies that $\frac{dD}{d\theta}<0$, since by our construction D = C - I. We also can see from the above that $\frac{de}{d\theta}>0$. We can now state our proposition.

Proposition 6

Under the conditions represented by our assumptions, for a particular level of cash availability, as the manager's productivity increases, the dividend decreases and the effort put in by the manager increases with respect to θ.

<u>Case 2:</u> $X(e,I,\theta)=\theta e\ln(I),\ k(e)=\frac{1}{2}Me^{2},\ h'(\theta)\leq 0$

In this section we use the production function that we used in our two-agent model, but instead of only two types of agents, we consider a continuum of agent types.

We use the equilibrium results derived in the previous section and examine the impact on dividends. Recall that the equilibrium investment and effort are jointly determined by

$$\begin{aligned} X_I[\bullet] &= 1+\gamma_I(\bullet)k'(\bullet)h(\theta) \\ X_e[\bullet] &= k'(\bullet)\left[1+\gamma_e(\bullet)h(\theta)\right]+\gamma(\bullet)k''(\bullet)h(\theta) \end{aligned} \tag{9}$$

In this case,

$$\gamma(\bullet)=\frac{X_\theta(\bullet)}{X_e(\bullet)}=\frac{e\ln(I)}{\theta\ln(I)}=\frac{e}{\theta}$$

Therefore, from the equilibrium equations (9) we have

$$\begin{aligned} \frac{\theta e}{I} &= 1 \\ \theta\ln(I) &= M e\left[1+\frac{1}{\theta}h(\theta)\right]+\frac{e}{\theta}M h(\theta) \\ &= M e\left[1+\frac{2}{\theta}h(\theta)\right] \end{aligned} \tag{10}$$

Taking the total differential with respect to θ, we get

$$-\frac{\theta e}{I^2}\frac{dI}{d\theta}+\frac{\theta}{I}\frac{de}{d\theta}+\frac{e}{I}=0$$

$$\frac{\theta}{I}\frac{dI}{d\theta}-M\left[1+\frac{2}{\theta}h(\theta)\right]\frac{de}{d\theta}+2Me\left[\frac{h(\theta)}{\theta^2}-\frac{h'(\theta)}{\theta}\right]+\ln(I)=0$$

Solving the above for $\frac{dI}{d\theta}$ and $\frac{de}{d\theta}$, we get

$$\frac{\dfrac{dI}{d\theta}}{\dfrac{\theta}{I}\left\{2Me\left[\dfrac{h(\theta)}{\theta^2}-\dfrac{h'(\theta)}{\theta}\right]+\ln(I)\right\}+\dfrac{e}{I}M\left[1+\dfrac{2}{\theta}h(\theta)\right]}$$

$$=\frac{\dfrac{de}{d\theta}}{\dfrac{e}{I}\dfrac{\theta}{I}+\dfrac{\theta e}{I^2}\left\{2Me\left[\dfrac{h(\theta)}{\theta^2}-\dfrac{h'(\theta)}{\theta}\right]+\ln(I)\right\}}$$

$$=\frac{1}{\dfrac{\theta e}{I^2}M\left[1+\dfrac{2}{\theta}h(\theta)\right]-\dfrac{\theta}{I}\dfrac{\theta}{I}}$$

(11)

Using the relationships in (10) and simplifying the denominator of the third term, we see that the simplified expression for the denominator of the third term would be $\frac{\theta^2}{I^2}\left[\ln(I)-1\right]$.

We can see from (11) that

$$\left\{\frac{dI}{d\theta},\frac{de}{d\theta}\right\}>0 \Leftrightarrow \ln(I)>1.$$

Also note that

$$\frac{dI}{d\theta}>0 \Leftrightarrow \frac{dD}{d\theta}<0$$

The interesting result of this section is that with costly effort and difference in productivity, dividend is a decreasing function of agent type. This result thus extends the result obtained by Miller and Rock (1985) . In their paper, Miller and

Rock had found that dividend is a monotonically increasing signal of agent type. We summarize our discussion in the following proposition.[9]

Proposition 7
With costly effort and differences in agent productivity, the dividend becomes a decreasing function of agent type.

The result is intuitive. For costly effort, the agent would substitute effort with investment. Higher the type of an agent, higher is his/her marginal productivity. Since the marginal cost of effort is the same for all agents, therefore higher the type of the agent, higher would be the substitution of investment for effort. More investment in our setting means less dividend. We shall then have dividend as a decreasing function of agent type.[10]

---

[9]It is interesting to note that the production function used by Miller and Rock (1985) is $F(I) = a \ln (I + b)$, $a, b > 0$.

[10]In Chapter 2 (the two agent case) we have examined the situation when the asymmetric information is about the initial cash endowment. We find that if we try to examine the same scenario here then the result becomes indeterminate (i.e we cannot determine the way dividends will change with θ). This is because in this section we first find the sign of $\frac{dI}{d\theta}$ and then we find the sign of $\frac{dD}{d\theta}$ from the relation $I = C - D(\theta)$. If C becomes a function of θ (which will be the case if we want to characterize asymmetric information by different initial cash endowment), then we cannot conclude about the sign of $\frac{dD}{d\theta}$ from our knowledge about the sign of $\frac{dI}{d\theta}$.

Chapter 5

Pricing Implication and Design of Empirical Test

5.1 Introduction

We have developed our model of dividend policy by positing dividends as devices to screen agents of different productivity. Productivity of an agent is an unobservable variable. However, our uninformed principal learns about managerial productivity by the dividend decision for a given level of cash availability. Since dividend level is publicly observable, and level of cash availability is also known to a large degree of confidence (through the use of publicly available, audited financial statements), the market would also be able to deduce the productivity of managers by observing dividend decisions. Therefore, market prices would react to dividends. In this chapter we determine the pricing implication of our model, and we also design an empirical test for it. We additionally examine the existing empirical studies in the light of the empirical implications of our model.

5.2 Pricing Implication

5.2.1 The Notation and the Setting

The following notations are used.

C Cash available for investment

I Investment level chosen by agent.

D Dividend level selected by an Agent = C - I.

θ Productivity parameter for agent.

e Effort put in by the agent.

$\tilde{\omega}_j$ The wage function offered to an agent.

$\tilde{Y}$ The fully observable stochastic output.

$\tilde{Y} = \theta e \ln(I) + \tilde{\varepsilon}$ where $\tilde{\varepsilon}$ is random noise with zero mean

The subscript j is used when the agent type is binary. This subscript denotes the corresponding values of various variables. When the distribution of the agents accords with a continuous distribution, θ, the agent type, is used as a function argument to indicate the values of various variables.

P_{BD} The price of the stock of the company *before* dividend declaration.

P_{ADC} The *cum-dividend* price of the stock of the company *after* dividend declaration.

P_{ADX} The *ex-dividend* price of the stock of the company *after* dividend declaration.

It is assumed, without loss of generality, that the number of shares issued is one. The sequence of events is given in **Table IV**.

1	2	3	4
C is observed by all. Agent privately observes his own type. Market sets P_{BD}	Principal offers a menu of wage contracts.	Agent declares his/her type and selects from menu. Agent pays D, spends effort and invests the balance of C over D. Market sets P_{ADC}. After Ex-Dividend date market sets P_{ADX}	The final output is realised and observed by all. Principal receives the payoff and agent receives the wage.

Table IV: The Sequence of Events for Pricing Implication.

The risk free rate is normalized at zero.

5.2.2 Discussion

In our sequence of events we have implicitly assumed that the market does not observe the selection of the contract by the agent. There are two reasons for this. Firstly, the executive compensation data is typically obtained from the proxy statements and the annual reports and these statements give the compensation data for the past year and *not for the coming year.*. Secondly, the contract is often implicit. Consider the following quotation from Murphy (1999).

> "An executive's wealth is *explicitly* (and mechanically) tied to the principal's objective (creating shareholder wealth) through his holdings of stock, restricted stock, and stock options. In addition, CEO wealth is *implicitly* tied to stock-price performance through accounting-based bonuses (reflecting the correlation between

accounting returns and stock-price performance) and through year-to-year adjustments in salary levels, target bonuses, and option and restricted stock grant sizes." (p.30).

Other researchers have also made the assumption that the information on compensation contract is private and not available to the market. For example Persons (1994) assumes that "the compensation contract is the private information of the manager and the board" (pp.426-427). Persons also writes the following as an elucidation.

> "Within the model, the assumptions of private contracting and lack of commitment may seem strained-why not just make the contract public and require large penalties if the contract is changed? In reality, there are several factors that, in my view, make these assumptions reasonable. They include competitive considerations, nonpecuniary and unobservable compensation, the limited length of employment contracts, and the complexity of the environment in which firms operate. First, a manager's compensation plan may reflect strategic considerations; considerations the firm does not want its competitors to know. This argues against a public contract. Second, one should view a manager's compensation broadly. In addition to a periodic pay-check, compensation can be nonpecuniary (plush office, corporate jet, directing corporate donations as the manager desires) and even unobservable (autonomy, cooperativeness from the board)."(p 437)

5.2.3 Pricing When Agent Type is a Binary Distribution

In this section we consider the pricing of stock when the agent type is binary. For the purpose of this section therefore the agent is either of L(ow) type or H(igh) type. Let p be the Principal's and the market's prior probability of an agent being of type L.

$$\begin{aligned} P_{BD} = \quad & p\left[D_L + \theta_L e_L \ln(C - D_L) - \omega_L\right] \\ + \quad & (1-p)\left[D_H + \theta_H e_H \ln(C - D_H) - \omega_H\right] \end{aligned}$$

After the agent declares his type, dividend is declared and this dividend conveys the information to the market about the agent type.

$$\begin{aligned} P_{ADC}\left(D = D_L\right) &= \left[D_L + \theta_L e_L \ln(C - D_L) - \omega_L\right] \\ P_{ADC}\left(D = D_H\right) &= \left[D_H + \theta_H e_H \ln(C - D_H) - \omega_H\right] \end{aligned}$$

The ex-dividend price is given by

$$\begin{aligned} P_{ADX}\left(D = D_L\right) &= \theta_L e_L \ln(C - D_L) - \omega_L \\ P_{ADX}\left(D = D_H\right) &= \theta_H e_H \ln(C - D_H) - \omega_H \end{aligned}$$

5.2.4 Pricing When Agent Type is a Continuous Distribution

Let the agent type be distributed (in the belief system of the principal and the market) according to a continuous distribution having the density function f(θ). The support of the distribution is [θ_{Min}, θ_{Max}].

$$P_{BD} = \int_{\theta_{MIN}}^{\theta_{MAX}} \left[D(\theta) + \theta\, e(\theta) \ln\left(C - D(\theta)\right) - \omega(\theta)\right] f(\theta)\, d\theta$$

$$\begin{aligned} P_{ADC}\left(D(\theta)\right) &= D(\theta) + \theta\, e(\theta) \ln\left(C - D(\theta)\right) - \omega(\theta) \\ P_{ADX}\left(D(\theta)\right) &= \theta\, e(\theta) \ln\left(C - D(\theta)\right) - \omega(\theta) \end{aligned}$$

Return on Dividend Announcement

$$= \frac{P_{ADC} - P_{BD}}{P_{BD}} = \frac{D(\theta)}{P_{BD}} + \frac{\theta e(\theta) \ln\left(C - D(\theta)\right)}{P_{BD}} - \frac{\omega(\theta)}{P_{BD}} - 1$$

5.3 Empirical Implications

Our model explains dividend as the component of a screening contract set up by an uninformed principal. Whereas the signalling models typically assume that managers are either entrepreneur-managers or that they desire to maximize some weighted combination of the wealth of shareholders (both current and new), our model does not assume any such thing. Our model assumes that the manager wants to maximize his net wealth, and that the principal recognizes this and sets up a screening contract to utilize the skill of the agent in the productive enterprise. This model is also interesting because this deals simultaneously with *moral hazard* (because the effort exerted by the agent is not observable) and *hidden information* (because the skill level of the agent is not known to the principal). The empirical implications of this model are opposite to the implications of the signalling models and the free cash flow hypothesis.

*We find that, contrary to the findings of the dividend models based on the signalling paradigm, a larger dividend -**conditioned on cash availability** - is bad news. That is, for a given level of cash availability, the lower type manager declares a higher dividend than that declared by a manager with higher productivity.* The intuition is that given a particular level of free cash flow, the manager with a higher productivity would be induced to invest more than the manager with a lower productivity. As a result, the manager with the lower productivity will

disburse more cash as a dividend compared to the manager with higher productivity.

If we assume that the earnings information proxies for the information on cash availability, then we would expect that a rational market would observe the dividend decision and infer that a higher level of dividend, *conditioned on earnings,* is an indicator of poor managerial type. We would therefore expect that, *conditioned on earnings*, the abnormal return on dividend announcement would be *negatively related* to the level of dividend. If, instead of returns we consider the stock valuation, then we would conclude that, *ceteris paribus*, the higher the dividend payment, the lower would be the stock value.

It would be useful here to highlight the differences between the empirical implications of our theory and the empirical implications of the signalling theories of dividend (e.g., Bhattacharyya, 1979; Miller and Rock, 1985; Williams, 1988; John and Williams, 1985; Heinkel,1978; Ambarish, John, and Williams, 1987) as well as the hypotheses based on free cash flow paradigm as enunciated in Easterbrook (1984) and Jensen (1986) .

When we compare the signalling models with our model, we find that the signalling models typically characterize the informational asymmetry by bestowing the manager or the insider with information about some aspect of the future cash flow. The manager knows this information but cannot *influence the probability distribution of the cash flow.* In contrast, in our model, not only does the agent or the manager have private information about the future cash flow, he/she can *determine (albeit at a cost) the expected value of the cash flow.* **Table**

V summarizes the informational asymmetry in our model vis a vis the informational asymmetry in signalling models. **Table VI** summarizes the choices made by agent in our model vis a vis the choices of agents in signalling models.

We see that the insider in the signalling models could not *choose* what the expected value of cash flow would be. By contrast, in our model the agent can *exert costly effort and decide* what the expected value of cash flow would be. However, the exertion of costly private effort means that moral hazard issues become relevant. If moral hazard were the only issue involved, then the efficient solution would be to "sell" the company to the agent at a price such that the ex ante expected pay off for the agent is ω_0-the reservation wage. In our model that solution is not optimal because the rent captured by the high type would be excessive. Therefore we find that in equilibrium we have to accept an inefficient investment for the low type agent so as to limit the rent captured by the high type.

	Model	Type of Information Asymmetry
	Our Model	The agent has private information about his/her productivity for a given level of effort and investment. The agent also chooses the privately costly unobservable effort. There is therefore both hidden information about agent type and moral hazard because of unobservable privately costly effort.
Signaling Models	Heinkel (1978)	The insider knows the rate of return on the capital investment.
	Bhattacharyya (1979)	The insider knows the support of the distribution of the cash flow.
	Miller and Rock (1985)	The insider knows the realized value of the cash flow in the first period.
	John and Williams (1985)	The insider knows the present value of the cash flow.
	Ambarish, John, and Williams (1987)	The insider knows the present value of the assets in place and the present value of the opportunities to invest.
	Williams (1988)	The insider knows the present value of the firm's risky real assets.

Table V: Informational asymmetry in our model vis a vis signaling models

	Model	Choice of Agent
	Our Model	The agent chooses the level of dividend and the level of effort. The choice of dividends also concurrently determines the level of investments because the amount of available cash is taken to be fixed. Higher type managers pay *lesser* dividend and exerts *more* effort.
Signaling Models	Heinkel (1978)	The insider manager chooses level of investment. The higher type manager signals his/her type by choosing a level of investment below the first-best level and paying the difference between the first-best level and the actual investment as dividend. Higher type managers pay out a higher dividend.
	Bhattacharyya (1979)	Manager commits a particular level of dividend. The higher the type, the higher is the dividend declared.
	Miller and Rock (1985)	Manager declares a dividend. Dividend declared is a monotonically increasing function of managerial type.
	John and Williams (1985)	Manager declares a dividend and sells stock to finance investment. The manager with better information declares a higher dividend and receives higher prices for its stock.
	Ambarish, John, and Williams (1987)	Manager selects both dividend and investment and sells/repurchases stock. In equilibrium manager signals both with dividends and investment.
	Williams (1988)	Manager selects both dividend and investment and sells/repurchases stock. In equilibrium, dividend is an increasing function of managerial information.

Table VI: Choices made by agent in our model vis a vis signaling models

The type of the hidden information is also important. When we construct a model where the difference between agents is characterized by different cash endowments, and where the productivity of the agents per unit of investment and per unit of effort is the same, we find that the outcome of such a model is similar to the outcome of the signalling models — viz. the higher type of agent pays a greater dividend. We get a result different from the results of the signalling models when, in addition to the moral hazard, the agent types are characterized by different productivity per unit of investment and effort. Note that if we leave out effort from our production function, there is no moral hazard in the model, and even in this case we find that the agent with lower productivity declares a higher dividend (see Appendix II). To summarize, our results are driven by the hidden information about the productivity of agents per unit of effort and per unit of investment.[11] Under the signalling theories, higher firm value is signalled by higher dividends. Therefore, under the signalling paradigm, dividend increase should result in higher abnormal return. Also, *ceteris paribus*, the value of the firm should be an increasing function of dividend. Our theory opposingly states

---

[11]Our model has both hidden information and hidden action. We have shown that it is hidden information that drives our result. A question could be asked on the necessity of including hidden action in our model. When I had set up the model, I did it both with hidden information and hidden action and then did the analysis. After I got the results that I did, then the question arose as to what is the real reason behind the results that I got. Further investigation showed that it is hidden information that drives our result. To modify the model now would be equivalent to looking at the answer and then changing the question. The important point here is that our results are driven by hidden information and holds even in the presence of hidden action. Therefore I have kept the original formulation of the model.

that, *conditioned on cash availability* (i.e., for a given level of cash availability; or, if earnings are taken as proxy for cash availability, then for a given level of earnings), higher dividend is an indication of lower agent type and should result in a lower abnormal return and lower firm value *(ceteris paribus)*.

The free cash flow conjectures of Easterbrook (1984) and Jensen (1986) posit that a higher dividend is better, because a higher dividend removes free cash from the hands of the managers, and the managers then have less money to waste. Also according to these conjectures, the announcement of higher dividends would lead to a higher abnormal return. However, we should point out that our concept of free cash flow is different from the way Easterbrook and Jensen used it. Jensen considers free cash flow to be the cash remaining with managers after all the positive NPV opportunities have been exhausted. To quote Jensen, "(f)ree cash flow is cash flow in excess of that required to fund all projects that have positive net present values when discounted at the relevant cost of capital"(p.323). In our model C is the cash available before the investment decision. We maintain that whenever the existing empirical studies have been conducted in a manner so as to condition on earnings, results are mostly in line with the empirical prediction of our theory.

5.4 Review of Existing Empirical Studies on Dividend

5.4.1 Empirical Studies that do not Condition on Earnings or Cash Availability.

In this section, we review the empirical research on dividend policies that have not conditioned their design on the level of either earnings or cash

availability. We restrict our focus to only those studies which have attempted to test two of the hypotheses of dividend policy, viz., Signalling Hypothesis and Free Cash Flow Hypothesis. We do not consider empirical studies of the Dividend Clientele Hypothesis, because such studies are only concerned with the taxation of dividend vis-a-vis taxation of capital gains.[12] Our position is that the empirical studies which examined the Signalling Hypothesis and/or Free Cash Flow Hypothesis should have included the level of earnings/cash availability as an important control variable. In this section, we shall review the findings of those empirical studies which *did not* control for earnings/cash availability. In the next section, however, we shall review studies which *did* control for earnings/cash availability, and we shall see how their findings are consistent with our empirical hypothesis.

Lang and Litzenberger (1989) test the cash flow signalling hypothesis vis-a-vis the overinvestment hypothesis. If the overinvestment hypothesis is correct, then a change in dividend will have greater impact on the return of the firm with overinvestment than on the return of the firm with underinvestment. Tobin's Q is used to classify a firm as an overinvestor or underinvestor. Tobin's Q is the ratio of the market value of the firm's equity and debt to the replacement cost of

---

[12] Papers which examine the dividend clientele hypothesis empirically are Michaely (1991); Poterba and Summers (1984); Black and Scholes (1974); Litzenberger and Ramaswamy (1982); Lewellen et al. (1978); Poterba (1986); Peterson, Peterson, and Ang (1985); Lakonishok and Vermaelen (1986); Chaplinsky and Seyhun (1990); Chen, Grundy, and Stambaugh (1990); Litzenberger and Ramaswamy (1979); Litzenberger and Ramaswamy (1980). These studies did not need to accommodate cash availability/earning because the focus was on tax impact on dividends vis a vis capital gains.

its assets. A firm with Q<1 is an overinvestor; a firm with Q>1 is an underinvestor.

The sample consists of firms which have changed the level of dividends. The daily risk adjusted returns for these firms are computed as the excess of actual return over the product of beta and the return on the market. It is found that the average return on dividend announcement days is significantly greater for firms with Q<1. Other tests look at the instances of dividend increases and dividend decreases separately. The conclusion of higher return on dividend announcement day for firms with Q<1 is unchanged. This study therefore supports the overinvestment hypothesis. *The study never used earnings or level of cash availability as a control variable.*

Bernheim and Wantz (1995) test the dividend signalling hypothesis vis a vis agency hypothesis in a differential taxation regime. Under dividend signalling hypothesis, dividends are used to signal some information about the firm. Dividends impose a dissipative cost on the shareholders, because the tax rate on dividends is usually higher than the tax rate on Capital Gains. Therefore, under the dividend signalling hypothesis, the higher the tax rate , the stronger will be the impact of a dollar of dividend on the stock return. Thus, with dividends as a signalling device, the abnormal return on a stock will increase with dividend increases in a stiffer tax regime. The alternative hypothesis holds that dividends might be determined through a process of managerial utility maximization; in this case, the abnormal return will show an opposite behaviour. Bernheim and Wantz formulate a regression model and conclude that data supports the dividend

signalling hypothesis. *They do not include earnings or cash availability as a control variable in this regression.*

Christie (1994), however, finds evidence that is somewhat contradictory to both information signalling theory and agency contracting theory. He looks at the market price reaction in response to a dividend cut vis-a-vis dividend omissions. According to both signalling and agency theories, a monotonic relation exists between prices and dividends. Yet Christie finds that the price reaction vis-a-vis dividend reductions and dividend omissions shows a U-shaped pattern. For less than 20% reduction, prices fall by about 4.95%, while for reductions exceeding 60%, prices fall by about 8.78%. In contrast, for dividend omissions, prices fall by about 6.94%. The techniques employed are two-tailed t-tests and regression. He also examine whether future dividends and growth opportunity could explain this U-shaped behaviour. He does not find such explanation supported by data. *He also does not control for earnings or cash availability.*

Like Christie, Yoon and Starks (1995) find that their study also does not support the free cash flow hypothesis of dividend policy. The investment opportunity set is proxied by Tobin's Q. They discover that, in general, low Q firms (Q<1) have higher dividend yields and larger dividend changes as well as smaller size. In addition, they determine that dividend increases are associated with increased capital expenditure for both low Q and high Q firms. This result is inconsistent with free cash flow hypothesis, but it is consistent with the cash flow signalling hypothesis of dividend policy. The empirical relationship between dividend change and revision of analysts' forecast is of interest to Yoon and

Starks, as well. They find that only dividend decreases are associated with revision of analysts estimates and not dividend increases. *As in the other studies reviewed in this section, Yoon and Starks (1995) do not condition on the level of earnings or cash flow.*

Sant and Cowan (1994) ascertain that managers omit dividends when the earnings variance increases. Their finding is consistent with the information signalling role of dividends. According to this view, managers omit dividends when they have less confidence in their own ability to predict earnings. Sant and Cowan select 381 cases of dividend omissions. For comparative purposes they construct a control sample of dividend-paying firms, and assign a date randomly to each of these firms as the artificial date of dividend omission. Thus they have a sample of firms which actually omitted dividends, and they have a control sample with random dates assigned as the date of dividend omission. They find that the earnings variance increases between the pre omission period and the post omission period. Additionally, the weekly beta increases between the pre-omission and the post omission period. Moreover, the volatility of analysts' forecasts and the variance of actual earnings increase between the pre-omission and the post-omission period. Consequently, for dividend omission cases, the volatility of earnings, volatility of analysts' forecasts, and beta increase after dividend omission. Sant and Cowan's research design compares the variables of the dividend omission sample and the control firms. They also run a cross-sectional regression of abnormal announcement returns on various explanatory variables. This regression, however, does not establish that the earnings variance

is a significant explanatory variable. *We would like to point out that this regression does not have any earnings term as a control explanatory variable.*

Aharony and Swary (1980) investigate whether dividend announcements release information to the market in addition to the information released by the earnings announcement, using an event study methodology. The data is classified into those instances where the earnings announcement precedes the dividend announcement, and those where the earnings announcement follows the dividend announcement. About 87% of the sample consists of cases where there is no change in dividend. Abnormal return is measured by using a market model. The examination of cumulative abnormal return over various windows relative to the dividend announcement date and earnings announcement date shows that an announcement of dividend increase/decrease does release information to the capital market over and above that released by an earnings announcement. This study also does not condition on the level of earnings; instead, it conditions on the *timing of earnings announcement vis-a-vis the timing of dividend announcement.*

Manuel, Brooks, and Schadler (1993) adopt an interesting approach to determine whether dividends, or earnings, or both are used as vehicles of information. They examine the question of whether new issues are valued differently by investors, depending upon whether the issue announcements precede or follow earnings and dividend announcements. They look at straight debt and equity issues and classify each issue in a 2x2 matrix — one dimension of the matrix being earnings announcements preceding or following the issue announcements, and the other dimension of the matrix being dividends

announcements preceding or following the issue announcements. Thus they have four groups for equity issues and four for debt issues. They generate the abnormal returns for these groups by a market model and examine the significance of the difference in the abnormal returns between the different groups. For both equity and debt issues they find that the issue timing vis-a-vis dividend announcement explains most of the abnormal returns. Like Aharony and Swary (1980) , these three researchers also do not condition on the level of earnings; instead they conditioned on the *timing of earnings announcement vis-a-vis the timing of dividend announcement.*

5.4.2 Empirical Studies that Condition on Earnings or Cash availability.

5.4.2.1 Studies that are consistent with our empirical hypothesis

We shall discuss the following studies: Cole (1980), Downes and Heinkel (1982) , Divecha and Morse (1983) and Bajaj (1988). While reporting these results, we have retained the original symbols used by the researchers and provided full explanation of those symbols.

5.4.2.1.1 Findings of Cole (1980)

Cole (1980) had the following research objective:"Now our objective is to investigate whether a) regular quarterly cash dividend announcements transmit information, b) dividend information is conditional on current earnings, current capital market access and current dividend amounts"(p.57)

Cole used data from 45 utility companies and from 50 industrial companies. He runs the following cross-sectional regression.

$$\tilde{e}_{j,0} = a_0 + a_1\tilde{d}_{j,0} + a_2\tilde{x}_{j,0} + a_3\tilde{v}_{j,0} + \left(a_4\tilde{x}_{j,0} + a_5\tilde{v}_{j,0}\right)\tilde{d}_{j,0} + a_5\left(\tilde{x}_{j,0}\tilde{v}_{j,0}\right) + \tilde{n}_{j,0}$$

where

$\tilde{e}_{j,0}$ = The residual return for stock j in the dividend announcement month. The residual return is computed by using a market model.

$\tilde{d}_{j,0}$ = The dividends per share

$\tilde{x}_{j,0}$ = The earnings per share

$\tilde{v}_{j,0}$ = The price per share lagged three months. Cole calls it a proxy for the access variable. It is meant to measure the "degree of access to external capital markets."

$\tilde{n}_{j,0}$ = The error term

The result obtained by Cole is reproduced in **Table VII** . We would like to point out that, consistent with our model, the coefficient of the dividend term (a_1) in the above regression is *negative and significant* for both the subsamples.

	Utility Subsample	Industrial Subsample
Number of Companies	45	50
Number of Dividend Announcements Used	1317	1061
a_0	0.051 (3.45)	0.0007 (0.96)
a_1	**-0.084 (-2.17)**	**-0.055 (-2.13)**
a_2	0.020 (1.95)	0.015 (1.58)
a_3	-0.002 (-3.88)	-0.0004 (-2.18)
a_4	0.006 (0.29)	-0.006 (-0.31)
a_5	0.004 (3.15)	0.0014 (2.67)
a_6	-0.0008 (-2.50)	-0.0001 (-2.57)
R^2	0.03	0.006
F ratio	8.10	1.9954
Critical F ratio at 5% significance	2.80	2.09

Table VII: Regression of residual announcement month returns for dividend announcements on various explanatory variables; t-statistics are given within parentheses.

Specifically, the following regression is run:

$$\tilde{e}_{j,0} = a_0 + a_1\tilde{d}_{j,0} + a_2\tilde{x}_{j,0} + a_3\tilde{v}_{j,0} + \left(a_4\tilde{x}_{j,0} + a_5\tilde{v}_{j,0}\right)\tilde{d}_{j,0} + a_5\left(\tilde{x}_{j,0}\tilde{v}_{j,0}\right) + \tilde{n}_{j,0}$$

where $\tilde{e}_{j,0}$ is the residual return for stock j in the dividend announcement month. The residual return is computed by using a market model; $\tilde{d}_{j,0}$ is the dividends per share; $\tilde{x}_{j,0}$ is the earnings per share; $\tilde{v}_{j,0}$ is the price per share lagged three months; $\tilde{n}_{j,0}$ is the error term. The results are from Cole, 1980 (from Tables 9 and 14, and pp. 77, 89, 94 and 101).

5.4.2.1.2 Findings of Downes and Heinkel (1982)

Downes and Heinkel (1982) test for the empirical validity of the Leland and Pyle (1977) model and the signalling model for dividends, using data from 297 Initial Public Offers (IPOs). The dividend signal valuation model is tested by setting up a regression linking the value of the firm with the product of earnings and a multiple. Thus Downes and Heinkel set up the following cross-sectional regression

$$V_j = m_j E_j + u_j$$

where

V_j	=	The total market value of equity after the initial offer.
E_j	=	The total normalized earning figure. The normalized earning was obtained from an exponential smoothing model using 10 years pre-issue data.
m_j	=	A multiple
u_j	=	An error term

The subscript j stands for the j-th firm. The multiple m_j is assumed to be a linear function of the following variables.

DEBT	The ratio of total debts to total assets after the equity issue.
IND	A dummy variable representing the industry classification of the issues. IND=1 for electronics manufacturing or computer services, otherwise IND=0
AGE	The log of the firm's age in years.
SLS	The log of the latest full years' sales.

GRW The continuously compounded rate of growth in sales over the most recent three-year period.

UWQ A dummy variable representing the prestige of the managing underwriter. UWQ=1 for prestigious underwriters and 0 otherwise.

HOT A dummy variable representing the "hot new issue market" that began during 1967 and continued till 1969. HOT=1 if the issue was offered on or after 1 July 1967 and 0 otherwise.

$\hat{\alpha}$ A measure of the Leland and Pyle (1977) signal of a firm's future cash flow. It is equal to $\alpha + \ln(1-\alpha)$, where α is the proportionate ownership retained by the entrepreneurs.

γ *A dividend policy dummy, which takes on a value of 1 if a dividend is declared in the prospectus, and zero otherwise.*

They estimated the above regression by OLS (Ordinary Least Squares)and found that the coefficient of the γ variable is negative (the estimated value was -2.91) and significant (t-value of -2.65).

In order to eliminate the impact of possible heteroscedasticity, Downes & Heinkel ran two other regressions. First, they ran a regression using the following regression equation:

$$V_j = m_j E_j + E_j e_j$$

The variable e_j is a disturbance term identically distributed across all firms. All other variables are as defined before. *This equation was estimated using WLS*

(Weighted Least Squares) procedure. They found that the coefficient of the γ variable is negative (the estimated value was -1.80) and significant at 10% level (t-value of -2.65).

Next, they ran another regression using the following regression equation

$$\ln\left(\frac{V_j}{E_j}\right) = \ln(m_j) + \ln(1 + e_j)$$

This equation was estimated using NLS (Non-linear Least Squares) procedure. They found that the coefficient of the γ variable is negative (the estimated value was -2.09) and significant at 5% level (t-value of -2.14). In a footnote to their paper, Downes & Heinkel state that they had tried payout ratio as another measure of dividend policy, and had also tried other adjustments to reduce heteroscedasticity. Still, the results were similar. *Therefore, we maintain that this finding is consistent with the empirical implication of our model, viz, conditioned on earnings, higher the dividend lower is the managerial productivity.*

5.4.2.1.3 Findings of Divecha and Morse (1983)

Divecha and Morse did an event study on dividend announcement. They take a sample of 1039 dividend increases for 735 firms from May 1977 through February 1979. They divide their sample into two parts — those where dividend payout has increased (Sample Size=668) and those where dividend payout has decreased (Sample Size=371). *They find that the excess return for the firms whose payout ratio have decreased is more than that for the firms whose payout ratio has increased.* Their plot of the Cumulative Average Residual (CAR) for days

surrounding the dividend announcement is shown in Figure 7. The CAR graph shows that the CAR for firms that had decreased the payout ratio is above the CAR for firms that had increased the payout ratio. Therefore the market perceives that the news for firms which have decreased payout ratio is better than the news for firms which have increased the payout ratio. *This is again consistent with our empirical testable prediction that given a particular level of earning, a higher dividend signifies relatively poor managerial type.*

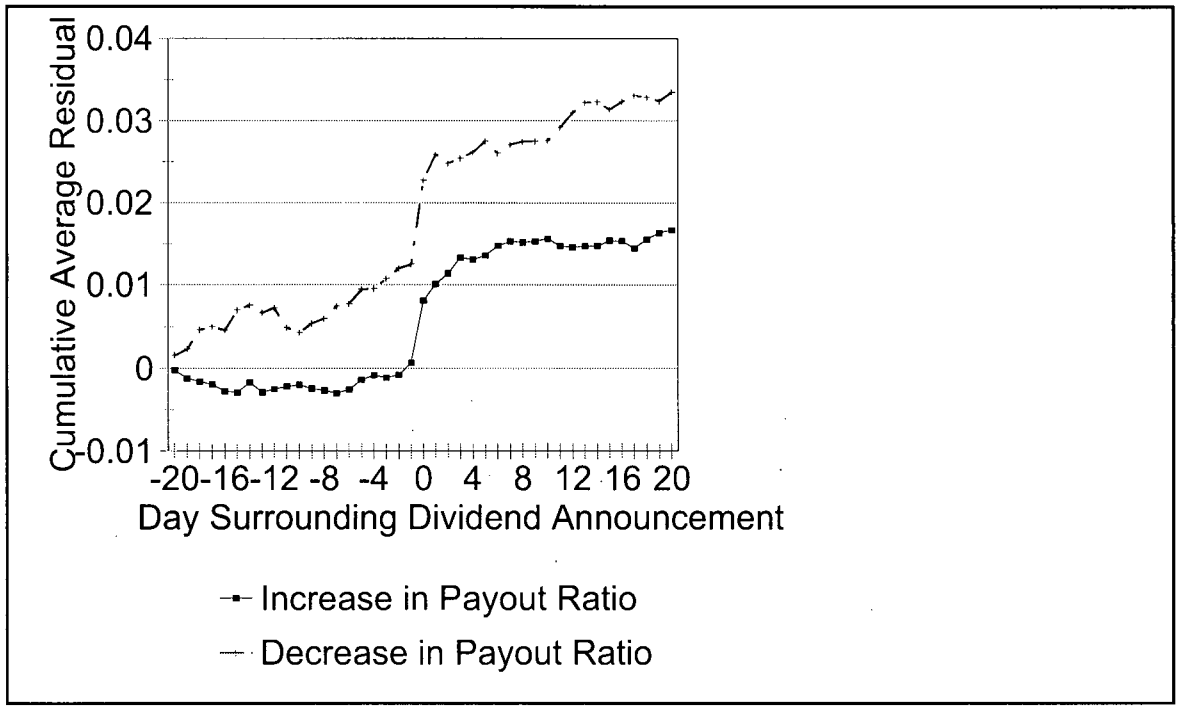

Figure 7: Cumulative Average Residuals for Days Surrounding an Announcement of an Increase in Dividends with Portfolios Partitioned on the Basis of an Increase or Decrease in the Payout Ratio. (Figure drawn from the data given in Table 3 of Divecha & Morse, 1983. A similar figure appears in Figure 3 of their paper.)

5.4.2.1.4 Findings of Bajaj (1988)

Bajaj selects a sample of firms from the Value Line Database II, based upon the following criteria :

(1) These firms do not belong to certain regulated industries.

(2) Quarterly earnings per share and dividends per share data was available from 1977-84. The firms must have paid uninterrupted, non zero quarterly dividends for this period.

(3) Quarterly dividend announcement dates are available on CRSP data and quarterly earnings announcement dates are available on COMPUSTAT.

(4) Daily return data is available on the CRSP daily return file.

(5) Fiscal year ends December 31. (Bajaj, 1988, p.36)]

This selection results in a sample of 252 firms. He then examines the dividend announcement dates and the earning announcement dates. Only those cases of dividend announcement and earnings announcement are selected where "the dividend announcement followed the earnings announcement by at least ten and at most twenty-five trading days" (Bajaj, 1988, p.37). Bajaj undertakes an event study on the earnings announcement days as well as on the dividend announcement days. The expected earning is measured by the "last forecast made [by the Value Line Investment Survey] before the earnings announcement" [p.38]. He classifies his results into the following four groups:

1. Those where the **actual earnings were higher than the expected earnings** by at least 10% and the **dividends have not changed from last quarter**.
2. Those where the **actual earnings were lower than the expected earnings** by at least 10% and the **dividends have not changed from last quarter**.
3. Those where the **actual earnings were higher than the expected earnings** by at least 10% and the **dividends were also increased by at least 10%** from last quarter.
4. Those where the **actual earnings were lower than the expected earnings** by at least 10% and the **dividends were also increased by at least 10%** from last quarter.

Bajaj's results are reproduced in **Table VIII**, **Table IX**, **Table X** and **Table XI**.[13]

Table VIII demonstrates that when earnings were higher than the expected earnings by at least 10%, *three* of the dividend announcement residual returns are positive and significant. As opposed to this, from **Table X**, we can see that when dividends are increased from their earlier level by at least 10% (the earnings in this case too were at least 10% more than the expected earnings), we only have *one* significant dividend announcement residual return. Looking at the total magnitude of these positive dividend announcement residual returns we see that, when dividends were *not* increased, the total significant dividend announcement residual return is 0.5892 (=0.1767+0.2545+0.1580; see **Table VIII**). *When*

[13]Bajaj's results are given in Tables 6,7,8 and 9 (pp.48-51] of his thesis (Bajaj, 1988).

dividends are increased, the corresponding significant dividend announcement residual return is 0.4353. **Table IX** and **Table XI** deal with the case where the actual earnings were lower than the expected earnings by at least 10%. Here also, we find that, when dividends are *not* increased, there is one significant *positive* dividend announcement residual return and its magnitude is 0.2286 (see **Table IX**). As opposed to this, when dividends are increased, we have two significant *negative* dividend announcement residual returns with a magnitude of - 1.2063 (= -0.6979 - 0.5084) (see **Table XI**). *We therefore reiterate that this finding is consistent with our empirical hypothesis.*

Cases where the **actual earnings were higher than the expected earnings** by at least 10% and the **dividends did not change** from last quarter. Sample Size=276		
Day	Earnings Announcement Residual Return (t-statistics in parentheses)	Dividend Announcement Residual Return (t-statistics in parentheses)
-5	-0.0192 (-0.1908)	-0.1223 (-1.3611)
-4	0.0936 (0.9986)	0.0041(0.0330)
-3	-0.0163 (-0.1381)	-0.1729 (-1.4661)
-2	**0.2751 (2.6408)**	0.0260 (0.2513)
-1	**0.6643 (5.1405)**	-0.0065 (-0.0568)
0	**0.3493 (2.6148)**	0.0936 (0.7052)
1	-0.0595 (-0.4732)	0.0357 (0.3544)
2	-0.0590 (-0.5656)	**0.1767 (1.9292)**
3	-0.0038 (-0.0393)	**0.2545 (2.2093)**
4	-0.0752 (-0.6493)	0.0208 (0.2057)
5	0.0511 (0.4758)	**0.1580 (1.7099)**

Table VIII: Results from Table 6 of Bajaj (1988). The bold entries are significantly different from zero at a 5% significance level in a one-tailed t-test. The significance in the t-test is calculated by me and not by Bajaj.

Cases where the **actual earnings were lower than the expected earnings** by at least 10% and the **dividends did not change** from last quarter. Sample Size=288		
Day	Earnings Announcement Residual Return (t-statistics in parentheses)	Dividend Announcement Residual Return (t-statistics in parentheses)
-5	0.0371 (0.3650)	0.0826 (0.7422)
-4	-0.1350 (-1.0570)	0.0557 (0.4666)
-3	0.0114 (0.1168)	-0.1187 (-1.1037)
-2	**-0.2597 (-2.3049)**	0.0104 (0.0722)
-1	**-0.4213 (-2.7660)**	-0.0476 (-0.4123)
0	-0.1855 (-1.5167)	-0.1100 (-0.7629)
1	-0.0829 (-0.6517)	0.0962 (0.8283)
2	0.0248 (0.2228)	-0.1168 (-1.0206)
3	-0.0455 (-0.4095)	0.1647 (1.4519)
4	**-0.2567 (-2.0501)**	0.1057 (0.8106)
5	-0.0368 (-0.3343)	**0.2286 (1.8677)**

Table IX: Results from Table 7 of Bajaj (1988). The bold entries are significantly different from zero at a 5% significance level in a one-tailed t-test. The significance in the t-test is calculated by me and not by Bajaj.

Cases where the **actual earnings were higher than the expected earnings** by at least 10% and the **dividends increased by at least 10%** from last quarter. Sample Size=46		
Day	Earnings Announcement Residual Return (t-statistics in parentheses)	Dividend Announcement Residual Return (t-statistics in parentheses)
-5	-0.3025 (-1.2663)	**0.4353 (1.7785)**
-4	0.0175 (0.0647)	-0.1011 (-0.3439)
-3	-0.4892 (-1.5985)	-0.1294 (-0.5755)
-2	0.3503 (1.0012)	0.0895 (0.3747)
-1	**0.4660 (1.8843)**	-0.0421 (-0.1545)
0	0.1617 (0.4946)	0.4773 (1.1864)
1	0.5249 (1.2037)	-0.1961 (-0.8146)
2	-0.1301 (-0.4350)	0.3049 (1.4369)
3	-0.1016 (-0.4144)	0.1069 (0.3845)
4	-0.1217 (-0.5471)	-0.1713 (-0.7022)
5	0.2744 (1.2288)	0.0351 (0.1573)

Table X: Results from Table 8 of Bajaj (1988). The bold entries are significantly different from zero at a 5% significance level in a one-tailed t-test. The significance in the t-test is calculated by me and not by Bajaj.

Cases where the **actual earnings were lower than the expected earnings** by at least 10% and the **dividends increased by at least 10%** from last quarter. Sample Size=28		
Day	Earnings Announcement Residual Return (t-statistics in parentheses)	Dividend Announcement Residual Return (t-statistics in parentheses)
-5	-0.2635 (-0.9913)	0.0011 (0.0032)
-4	-0.2231 (-0.6015)	0.2251 (0.7661)
-3	0.1607 (0.5557)	**-0.6979 (-2.1958)**
-2	0.1281 (0.3243)	-0.3320 (-0.9582)
-1	-0.7411 (-0.8482)	0.1388 (0.5195)
0	-0.2634 (-0.7073)	**-0.5084 (-1.7383)**
1	0.2290 (0.7393)	0.5148 (1.5516)
2	0.3250 (0.9837)	0.0254 (0.0990)
3	**-0.7677 (-2.4611)**	0.0230 (0.1040)
4	0.6385 (1.1327)	0.1909 (0.4063)
5	**-0.7680 (-2.1386)**	0.2552 (0.7803)

Table XI: Results from Table 9 of Bajaj (1988). The bold entries are significantly different from zero at a 5% significance level in a one-tailed t-test. The significance in the t-test is calculated by me and not by Bajaj.

5.4.2.2 Studies that are not consistent with our empirical hypothesis

We have discussed empirical studies that have controlled for earnings while examining the impact of dividends. We have shown that the findings of these studies are consistent with our empirical hypothesis. However, during our literature research, we have found two papers in which the researchers control for earnings, but obtain findings contrary to what our empirical hypothesis would suggest. These studies are Pettit (1972) and Denis, Denis, and Sarin (1994).

Pettit (1972) investigates the impact of dividend announcement on returns vis-a-vis earnings announcement. The unexpected returns in each period are computed according to the market model. For each month of the study, the dividend announcements are examined and are classified into seven categories. Petit then further classifies the reported earnings according to whether they decrease or increase from the expected earnings. An event study approach is adopted for each of the earnings categories and for each of the dividend categories. The graphs of the cumulated abnormal returns look very similar for both earnings increase and decrease. The event study is conducted separately for the earnings increase and for the earnings decrease cases where the event dates are the dividend announcement dates. Pettit concludes that dividends give more information than earnings.

Denis, Denis, and Sarin (1994) examine the three hypotheses of dividend policy, namely, the signalling hypothesis, the free cash flow hypothesis, and the dividend clientele hypothesis jointly. They, *inter alia*, regress a two day announcement period return on the following variables:

CHNG	=	The change in dividend, divided by the market value of equity two days prior to dividend announcement.
YLD	=	The most recent quarterly dividend, divided by the market value of equity two days prior to dividend announcement
CFLOW	=	Operating income before depreciation minus interest expenses, taxes, preferred dividends, and common dividends, all divided by total assets.
QDUM	=	A dummy variable which takes a value of 1 if Tobin's Q is greater than 1, and a value of 0 otherwise.
Q x CFLOW	=	Product of Tobin's Q and CFLOW.

Furthermore, they find that the coefficient of the CFLOW term is *negative and significant* (the coefficient was -2.80 with a t-ratio of -2.06), whereas the coefficient of the CHNG term is positive, large, and significant (the coefficient was 419.27 with a t-ratio of 24.40).

5.4.3 Discussion

We have seen that in a number of cases of empirical investigation of dividends, whenever the research design controls for earnings, the impact of dividend announcement concurs with our empirical hypothesis. Dividend announcements are usually accompanied with earnings announcements. Domian (1987) finds that earnings "cause" dividends in the sense of Granger causality. Healy and Palepu (1988) look at a sample of dividend initiations (i.e., first time dividend payment in a long time) and at another sample of dividend omissions.

It is found that dividend initiating firms have significant positive change in earnings in the year of dividend change and the year prior to it. This positive change in earning continues for two years after the dividend initiation. For the dividend omission sample, the significant negative change occurs in the year of dividend omission and the year before it. However, the earnings change subsequent to the dividend omission is positive and significant for two years.

The findings by Domian (1987) and Healy and Palepu (1988) underscore the importance of controlling for the earnings variable in an empirical investigation of the impact of dividends. Our theory was developed by assuming that the amount of cash available is known, and that wage contracts are used to entice managers to declare dividends (thereby revealing the quantum of investments) in such a manner that the most efficient manager gets a first best efficient contract and the other managers get less efficient contracts. It was shown earlier that this scenario results in a situation where the dividend level, *conditional on cash availability,* is inversely related to the managerial type. We would like to point out that cash availability can be deduced from the earnings statement (assuming, of course, that the earnings statement is properly prepared and that no earnings management is taking place). The Granger causality of earnings to dividends as found by Domian is thus in tune with our assumption of fully observable cash flow. The findings of Healy and Palepu reinforce the validity of our assumption.

5.5 Empirical Design for Testing Our Model

We shall now develop the research design for testing our model empirically. We have developed the following relation:

Return on Dividend Announcement

$$= \frac{P_{ADC} - P_{BD}}{P_{BD}} = \frac{D(\theta)}{P_{BD}} + \frac{\theta e(\theta) \ln(C - D(\theta))}{P_{BD}} - \frac{\omega(\theta)}{P_{BD}} - 1 \tag{12}$$

where

C Earnings or cash available for investment

I Investment level chosen by agent.

D Dividend level selected by an Agent = C-I

θ Productivity parameter for agent.

e Effort put in by the agent.

ω The wage function offered to an agent.

P_{BD} The price of the stock of the company *before* dividend declaration.

P_{ADC} The *cum-dividend* price of the stock of the company *after* dividend declaration.

5.5.1 Procedure for Empirical Testing

The model as enunciated in (12) can be stated in the following cross-sectional regression equation form.

$$\tilde{r}_j = \beta_0 + \beta_1 \frac{D_j}{P_j} + \beta_2 \frac{\ln(C_j - D_j)}{P_j} + \beta_3 \frac{\omega_j}{P_j} + \tilde{\varepsilon}_j \tag{13}$$

where

C_j Earnings or cash available for firm j

D_j Total dividend level for firm j = Dividend per share times the number of stock outstanding.

ω_j The compensation paid to top-executives of firm j.

P_j Total equity capitalization for the firm = The price of the stock of the company *before* dividend declaration times the number of stock outstanding.

$\tilde{\varepsilon}_j$ The random error term for firm j.

β s The regression coefficients.

The regression coefficients in (13) cannot be used to test the validity of (12), because if (12) were true then β_2 would be equal to θe, i.e the coefficient β_2 would be different for different combinations of C_j and D_j. However, the regression procedures used for estimating the coefficients assume that these coefficients are constant. Note that we cannot get away from this problem by including θ and e explicitly in our regression formulation, because both θ and e are unobservable variables. Since the coefficient β_2 depends on the level of C_j and D_j, we can try to test our model using a dummy variable regression procedure. In order to use the dummy variable procedure, however, we need to discretize the levels of C_j and D_j. Since our model predicts that, conditioned on the level of earnings or the cash available, higher dividend means a lower managerial productivity and lower managerial effort, we propose that the discretization be done with respect to the payout ratio, i.e., the ratio of D_j to C_j. We therefore suggest the following empirical procedure for testing our model:

Step1: Collect the accounting data, price data and compensation data for firms with a continuous dividend record.

Step 2: Sort the firms according to the pay-out ratios and classify them into deciles.

Step 3: Do an event study on the firms on dividend announcement day and compute the abnormal return on a three-day window around the dividend announcement day. This will give us r_j.

Step 5: Run the following regression:

$$\begin{aligned}\tilde{r}_j = \quad & \beta_0 + \beta_1 \frac{D_j}{P_j} + \beta_2 \frac{\ln\left(C_j - D_j\right)}{P_j} + \beta_3 \frac{\omega_j}{P_j} \\ & + \sum_{k=2}^{10} \gamma_k \, \mathrm{DUMMY}_k \frac{\ln\left(C_j - D_j\right)}{P_j} + \tilde{\varepsilon}_j\end{aligned} \tag{14}$$

where

DUMMY_k is a dummy variable which takes a value 1 if firm j belongs to the k-th decile and zero otherwise.

γ_k Coefficients of the dummy terms.

Other variables are as defined before.

Step 6: If our theory is correct then we should observe that
$0 > \gamma_2 > \gamma_3 > \gamma_4 > \gamma_5 > \gamma_6 > \gamma_7 > \gamma_8 > \gamma_9 > \gamma_{10}$.

Test to see if our estimated coefficients satisfy the above criterion. A rank test may be used.

5.5.2 Discussion

In the test proposed in the previous section, we have not included control variables. This was done for reasons of brevity and to describe the rationale of the procedure comprehensively. While performing the actual empirical tests, we shall

have to control for various plausible influential factors in order to ensure that our results are rigorous. The two control variables that we propose to include in estimating the parameters in (14) are size and debt-equity ratio. We should also control for the price effects due to unexpected earnings. We can increase the robustness of our results by estimating (14) separately for different industries. However, this approach will depend on the quantum of data available. We would like to note, in addition, that the discretization of data based on payout ratio is inspired by the discretization used in Litzenberger and Ramaswamy (1982) and Litzenberger and Ramaswamy (1980) . There the researchers assume that "individuals fall into five tax clienteles and each clientele holds one-fifth of the market value of all New York Stock Exchange (NYSE) stocks."[p.477]

Our model does not allow any external financing. In the real world however managers do have access to capital markets. We need to control for this in testing our model empirically. We could control for this by ensuring that in our sample no access to capital market takes place within some specified days of declaration of dividend. We also feel that inclusion to debt-equity ratio as a control variable also helps control for the presence of external financing. We can also look for the extent of external financing and include it as a control variable.

In developing our model we have relied on the empirical observation that executive compensation is related to dividend and output. In developing our empirical test however we have used the level of executive compensation. This opens our proposed test to an errors in variables bias. One way to obviate this problem is to estimate the compensation function (assuming the functional

relationship is linear) and then use the estimated compensation function in place of ω_j in the testing procedure outlined earlier.

We can also perform a more robust non-parametric test by noting that (12) can be written as

$$\text{Return on Dividend Announcement-}\frac{D(\theta)}{P_{BD}} + \frac{\omega(\theta)}{P_{BD}} + 1 = \frac{\theta e(\theta)\ln\left(C - D(\theta)\right)}{P_{BD}}$$

We can see from the above equation that if our theory is correct, then the quantity on the left hand side must be a decreasing function of pay-out ratio. We can easily check this by means of a rank test. Finally, we can also try using non-linear estimation techniques to estimate the parameters in (13).

Chapter 6

Conclusion

We have used the Principal-Agent paradigm to explain dividends, and we find that, when the hidden information is about the productivity of agent, then contrary to the prediction of the dividend theories based on signalling paradigm, our model posits dividends *conditioned on earnings* to be unfavourable news. That is, for a given level of earnings, the lower type manager declares a higher dividend than that declared by a higher type manager. This is intuitive, because a higher type manager has higher productivity, and, therefore, it makes sense to induce the higher type manager to make higher investment and thus pay less dividend. We also find that costly private effort and difference in productivity can flip the relationship between dividend and managerial type from being monotone-increasing (a la Miller and Rock, 1985) to being monotone-decreasing. The signalling models, like Miller and Rock's and Heinkel's (1978) have underinvestment as a cost, and the higher type manager underinvests. In contrast, in our model the higher type manager achieves the first best investment level and the lower type manager underinvests so as to limit the information rent paid to the higher type manager.

We find that the nature of hidden information is important. When we characterize the asymmetric information as the information about initial cash

endowment, we find results similar to the signaling theories and free cash flow hypothesis-namely that higher agent type results in higher dividends.

The proof of the pudding is in eating, and the validity of a theory is the degree of its congruence with empirical reality. We have designed an empirical test for testing our model. Future research efforts shall be directed towards actually conducting this empirical test, and towards extending this model to a multi-period set up.

Bibliography

1. Aharony, Joseph, and Itzhak Swary. 1980. Quarterly Dividend and Earnings Announcements and Stockholders' Returns: An Empirical Analysis. *Journal Of Finance* 35, no. 1: 1-12.

2. Allen, Franklin, and Roni Michaely. 1995. Dividend Policy. *Finance.* editor Robert A. Jarrow, Vojislav Maksimovic, and W. T. Ziemba. Handbooks in Operations Research and Management Science, 9. Amsterdam ; New York: Elsevier.

3. Ambarish, Ramasastry, Kose John, and Joseph Williams. 1987. Efficient Signalling with Dividends and Investments. *Journal Of Finance* XLII, no. 2: 321-43.

4. Bajaj, Mukesh. 1988. "Rational dividend policy: theory and evidence." Phd, University of California, Berkeley (0028).

5. Banker, Rajiv D., Somnath Das, and Srikant M. Datar. 1993. Complementarity of Prior Accounting Information:The Case of Stock Dividend Announcement. *The Accounting Review* 68, no. 1: pp 28-47.

6. Bar-Yosef, Sasson, and Itzhak Venezia. 1991. Earnings Information and the Determination of Dividend Policy. *Journal of Economics and Business* 43, no. 3: 197-214.

7. Bernheim, B. Douglas, and Adam Wantz. 1995. A Tax-Based Test of the Dividend Signaling Hypothesis . *American Economic Review* 85, no. 3: 532-51.

8. Bhattacharyya, Sudipto. 1979. Imperfect Information, dividend policy, and "the bird in the hand" fallacy. *Bell Journal of Economics* 10: pp 259-70.

9. Black, Fischer. 1976 . The Dividend Puzzle. *Journal-of-Portfolio-Management* 2, no. 2: 5-8.

10. Black, Fischer, and Myron Scholes. 1974. The Effects of Dividend Yield and Dividend Policy On Common Stock Prices and Returns. *Journal of Financial Economics* 1: 1-22.

11. Bond, Michael T., and Mbodja Mougoue. 1991. Corporate Dividend Policy and the Partial Adjustment Model. *Journal of Economics and Business* 43, no. 2: pp 165-77.

12. Boyd, John H., and Ravi Jagannathan. 1994. Ex-Dividend Price Behaviour of Common Stocks. *Review of Financial Studies* 7, no. 4: 711-41.

13. Brealey, Richard A., Stewart C. Myers, Alan J.Marcus. 1999. Fundamentals of Corporate Finance. 2nd ed.Irwin McGraw-Hill Limited.

14. Brennan, Michael J., and Anjan V. Thakor. 1990. Shareholder Preferences and Dividend Policy. *Journal Of Finance* 45, no. 4: 993-1018.

15. Chang, S. J., and Son-Nan Chen. 1991. Information Effects of Earnings and Dividend Announcements on Common Stock Returns: Are They Interactive. *Journal of Economics and Business* 43, no. 2: 179-92.

16. Chaplinsky, Susan, and H. Nejat Seyhun. 1990. Dividends and Taxes:Evidence on Tax-Reduction Strategies. *Journal of Business* 63, no. 2: 239-60.

17. Chen, Nai-Fu, Bruce Grundy, and Robert F. Stambaugh. 1990. Changing Risk, Changing Risk-Premiums, and Dividend Yield Effects. *Journal of Business* 63, no. 1,part 2: S51-S70.

18. Christie, William G. 1994. Are Dividend Omissions Truly the Cruelest Cut of All. *Journal of Financial and Quantitative Analysis* 29, no. 3: 459-80.

19. Cole, John Anthony. 1980. "Theory and Evidence Regarding the Information Content in Dividends." The University Of Michigan.

20. Crystal, Graef S. 28 August 1989. Incentive Pay that doesn't work. *Fortune,* pp. 101-4.

21. Datta, Sudip, and Upinder S. Dhillon. 1993. Bond and Stock Market Response to Unexpected Earnings Announcements. *Journal of Financial and Quantitative Analysis* 28, no. 4: 565-77.

22. DeAngelo, Harry, and Linda DeAngelo. 1990. Dividend Policy and Financial Distress: An Empirical Investigation of Troubled NYSE Firms. *Journal Of Finance* 45, no. 5: 1415-31.

23. ———. 1991. Union negotiations and corporate policy. *Journal of Financial Economics* 30: pp 3-43.

24. DeAngelo, Harry, Linda DeAngelo, and Douglas J. Skinner. 1992. Dividends and Losses. *Journal Of Finance* 47, no. 5: 1837-63.

25. ———. 1996. Reversal of fortune-Dividend signalling and the disappearance of sustained earnings growth. *Journal of Financial Economics* 40: pp 341-71.

26. Denis, David J., Diane K. Denis, and Atulya Sarin. 1994. The Information Content of Dividend Changes:Cash Flow Signaling,Overinvestment,and Dividend Clienteles. *Journal of Financial and Quantitative Analysis* 29, no. 4: 567-87.

27. Dhillon, Upinder S., and Herb Johnson. 1994. The Effect of Dividend Changes on Stock and Bond Prices. *Journal Of Finance* 49, no. 1: 281-89.

28. Divecha, Arjun, and Dale Morse. June1983. Market Responses to Dividend Increases and Changes in Payout Ratios . *Journal-of-Financial-and-Quantitative-Analysis* 18, no. 2: 163-73.

29. Domian, Dale Louis. 1987. "Relations among dividends, earnings, and investment." Phd, University of Minnesota (0130).

30. Downes, David H., and Robert Heinkel. 1982. Signalling and the Valuation of Unseasoned New Issues. *Journal Of Finance* 37, no. 1: 1-10.

31. Easterbrook, Frank H. 1984. Two Agency-Cost Explanations of Dividends. *American Economic Review* 74, no. 4: 650-659.

32. Faynzilberg, Peter S, and Kumar, Praveen. 1997. Optimal Contracting of Separable Production Technologies. *Games and Economic Behavior* 21, no.1-2, pp 15-39.

33. Ferreira White, Lourdes. 1996. Executive Compensation and Dividend Policy. *Journal-of-Corporate-Finance* 2, no. 4: 335-58.

34. Fudenberg, Drew, and Jean Tirole. 1991. Game Theory. Massachusetts Institute of Technology.

35. Gaver, Jennifer J., and Kenneth M. Gaver. 1993. Additional evidence on the association between the investment opportunity set and corporate financing, dividend and compensation policies. *Journal of Accounting and Economics* 16: pp 125-60.

36. Goetzmann, Willam N., and Philippe Jorion. 1993. Testing the Predictive Power of Dividend Yields. *Journal Of Finance* 48, no. 2: 663-79.

37. Healy, Paul M. 1985. The Effect of Bonus Schemes on Accounting Decisions. *Journal of Accounting and Economics* 7: 85-107.

38. Healy, Paul M., and Krishna G. Palepu. 1988. Earnings Information Conveyed By Dividend Initiations And Omissions. *Journal of Financial Economics* 21: pp 149-75.

39. ———. 1990. Effectiveness of Accounting-Based Dividend Covenants. *Journal of Accounting and Economics* 12: pp 97-123.

40. Heinkel, Robert Lee. 1978. "Essays on Financial Markets with Asymmetric Information." Microfiche, University of California, Berkeley.

41. Jensen, Gerald R., Donald P. Solberg, and Thomas S. Zorn. 1992. Simultaneous Determination of Insider Ownership, Debt and Dividend Policies. *Journal of Financial and Quantitative Analysis* 27, no. 2: 247-63.

42. Jensen, Michael C. 1986. Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review* 76, no. 2: 323-29.

43. Jensen, Michael C., and Kevin J. Murphy. 1990. Performance Pay and Top-Management Incentives. *Journal of Political Economy* 98, no. 2: 225-64.

44. John, Kose, and Joseph Williams. 1985. Dividends, dilution and taxes: A Signalling Equilibrium. *Journal Of Finance* 40, no. 4: 1053-70.

45. Kao, Chihwa, and Chunchi Wu. 1994. Tests of Dividend Signalling Using the Marsh-Merton Model: A Generalised Friction Approach. *Journal of Business* 67, no. 1: 45-68.

46. Kaplan, Steven N., and David Reishus. 1990. Outside directorships and corporate performance. *Journal of Financial Economics* 27: 389-410.

47. Kumar, Praveen. 1988. Shareholder-Manager Conflict and the Information Content of Dividends. *Review of Financial Studies* 1, no. 2: pp 111-36.

48. Laffont, Jean-Jacques, and Jean Tirole. 1993. *A Theory of Incentives in Procurement and Regulation.* Cambridge, Massachusettes: MIT Press.

49. Lakonishok, Josef, and Theo Vermaelen. 1986. Tax-Induced Trading Around Ex-Dividend Days. *Journal of Financial Economics* 16: 287-319.

50. Lambert, Richard A., William N. Lanen, and David F. Larcker. 1989. Executive Stock Option Plans and Corporate Dividend Policy. *Journal of Financial and Quantitative Analysis* 24, no. 4: 409-25.

51. Lang, Larry H. P., and Robert H. Litzenberger. 1989. Dividend Announcements - Cash Flow Signalling vs. Cash Flow Hypothesis? *Journal of Financial Economics* 24: 181-91.

52. Larcker, David F. 1983. The Association between Performance Plan Adoption and Corporate Capital Investment. *Journal of Accounting and Economics* 5, no. 1: 3-30.

53. Lee, Charles M. C. 1996. Measuring Wealth. *CA Magazine* , no. April 1996: 32-34.

54. Leland, Hayne E., and David H. Pyle. 1977. Informational Asymmetries, Financial Structure, and financial Intermediation. *Journal Of Finance* 32, no. 2: pp 371-87.

55. Lewellen, Wilbur G., Claudio Loderer, and Kenneth Martin. 1987. Executive Compensation and Executive Incentive Problems: An Empirical Analysis. *Journal of Accounting and Economics* 9, no. 3: 287-310.

56. Lewellen, Wilbur G., Kenneth L. Stanley, Ronald C. Lease, and Gary G. Schlarbaum. 1978. Some Direct Evidence on the Dividend Clientele Phenomenon. *Journal Of Finance* 33, no. 5: 1385-99.

57. Lintner, John. 1956. Distribution Of Incomes Of Corporations Among Dividends, Retained Earnings and Taxes. *American Economic Review* 46: 97-113.

58. Litzenberger, Robert H., and Krishna Ramaswamy. 1980. Dividends, Short Selling Restrictions, Tax-Induced Investor Clienteles and Market Equilibrium. *Journal of Finance* 35, no. 2: 469-82.

59. ———. 1979. The Effect of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence. *Journal of Financial Economics* 7, no. 2: 163-95.

60. ———. 1982. The Effects of Dividends on Common Stock Prices-Tax Effects or Information Effects? *Journal Of Finance* 37, no. 2: 429-43.

61. Long, Michael S., Ileen B. Malitz, and Stephan E. Sefcik. 1994. An Empirical Examination of Dividend Policy Following Debt Issues. *Journal of Financial and Quantitative Analysis* 29, no. 1: 131-44.

62. Manuel, Timothy A., LeRoy D. Brooks, and Frederick P. Schadler. 1993. Common Stock Price Effects of Security Issues Conditioned by

Current Earnings and Dividend Announcements. *Journal of Business* 66, no. 4: 571-93.

63. Michaely, Roni. 1991. Ex-Dividend Day Stock Price Behaviour: The Case of the 1986 Tax Reform Act. *Journal Of Finance* 46, no. 3: 845-59.

64. Miller, Merton H., and Franco Modigliani. 1961. Dividend Policy, Growth and the Valuation of Shares. *Journal of Business* 34, no. 4: pp 411-33.

65. Miller, Merton H., and Kevin Rock. 1985. Dividend Policy under Asymmetric Information. *Journal Of Finance* 40, no. 4: pp 1031-51.

66. Miller, Merton H., and Myron S. Scholes. 1978. Dividends and Taxes. *Journal of Financial Economics* 6: pp 333-64.

67. Murphy, K. J, 1999. Executive Compensation. Working Paper

68. Myers, Stewart C., and Nicholas S. Majluf. 1984. Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have. *Journal of Financial Economics* 13: 187-221.

69. Myerson, Roger B. 1982. Optimal Coordination Mechanisms in Generalized Principal-Agent Problems. *Journal of Mathematical Economics 10, no.1: pp 67-81.*

70. Ofek, Eli. 1993. Capital structure and firm response to poor performance-An empirical analysis. *Journal of Financial Economics* 34: pp 3-30.

71. Offer, Aharon R., and Daniel R. Siegel. 1987. Corporate Financial Policy, Information and Market Expectations: An Empirical Investigation of Dividends. *Journal Of Finance* 42, no. 4: 889-911.

72. Paik, Tae-Young, and Pradyot K. Sen. 1995. Project Evaluation and Control in Decentralized Firms: Is Capital Rationing Always Optimal? *Management Science* 41, no. 8: 1404-14.

73. Pavlik, Ellen L., Thomas W. Scott, and Peter Tiessen. 1993. Executive Compensation: Issues and Research. *Journal of Accounting Literature* 12: 131-89.

74. Penman, Stephen H. 1983. The Predictive Content of Earnings Forecasts and Dividends. *Journal Of Finance* 38, no. 4: 1181-99.

75. Persons, J. C. 1994. Renegotiation and the Impossibility of Optimal Investment. *Review-of-Financial-Studies*, 7, no. 2: 419-49.

76. Peterson, Pamela P., David R. Peterson, and James S. Ang. 1985. Direct Evidence On The Marginal Rate Of Taxation On Dividend Income. *Journal of Financial Economics* 14: pp 267-82.

77. Pettit, R. Richardson. 1972. Dividend Announcements, Security Performance and Capital Market Efficiency. *Journal Of Finance* 27, no. 5: 993-1007.

78. Poterba, James M. 1986. The Market Valuation of Cash Dividends - The Citizens Utilities Case Reconsidered. *Journal of Financial Economics* 15: 395-405.

79. Poterba, James M., and Lawrence H. Summers. 1984. New Evidence That Taxes Affect the Valuation of Dividends. *Journal Of Finance* 39, no. 5: 1397-415.

80. Rasmusen, Eric. *Games and Information* 2nd ed.Oxford:Blackwell Publishers Ltd.

81. Ravid, S. Abraham, and Oded H. Sarig. 1991. Financial Signalling by Committing to Cash Outflows. *Journal of Financial and Quantitative Analysis* 26, no. 2: 165-80.

82. Riley, John G. 1979. Informational Equilibrium. *Econometrica;* 47, no. 2: 331-59.

83. Sant, Rajiv, and Arnold R. Cowan. 1994. Do dividends signal earnings? The case of omitted dividends. *Journal of Banking and Finance* 18: 1113-33.

84. Shefrin, Hersh M., and Meir Statman. 1984. Explaining Investor Preference For Cash Dividends. *Journal of Financial Economics* 13: pp 253-82.

85. Smith, Clifford W. Jr., and Ross L. Watts. 1992. The investment opportunity set and corporate financing, dividend and compensation policies. *Journal of Financial Economics* 32: 263-92.

86. Tirole, Jean. 1988. The Theory of Industrial Organization. Massachusetts Institute of Technology.

87. Varian, Hal R. 1992. *Microeconomic Analysis.* 3rd ed. New York: W.W.Norton & Company.

88. Viswanath, P. V. 1993. Strategic Considerations, the Pecking Order Hypothesis, and Market Reactions to Equity Financing. *Journal of Financial and Quantitative Analysis* 28, no. 2: 213-34.

89. Williams, Joseph. 1988. Efficient Signalling with Dividends, Investment and Stock Repurchases. *Journal Of Finance* 43, no. 3: pp 737-47.

90. Wruck, Karen Hopper. 1994. Financial Policy, internal contrl and performance: Sealed Air Corporation's leveraged special dividend. *Journal of Financial Economics* 36: pp 157-92.

91. Yoon, Pyung Sig, and Laura T. Starks. 1995. Signalling, Investment Opportunities, and Dividend Announcements. *Review of Financial Studies* 8, no. 4: 995-1018.

Appendix I: Proofs of Propositions

Lemma A1.1

$$\frac{\partial D_j}{\partial \beta_D} > 0 \qquad \frac{\partial D_j}{\partial \beta_Y} < 0$$

$$\frac{\partial e_j}{\partial \beta_D} < 0 \qquad \frac{\partial e_j}{\partial \beta_Y} > 0$$

$$j = H, L$$

Proof:

From (IC′) we have

$$\beta_D(j) - \frac{\beta_Y(j)\theta_j e_j}{C - D_j} = 0 \tag{A1.1}$$

$$\beta_Y(j)\,\theta_j \ln(C - D_j) = M\,e_j \qquad j = L, H \tag{A1.2}$$

Totally differentiating **(A1.1)** with respect to β_D and β_Y, we get

$$\left[1 - \frac{\beta_Y(j)\,\theta_j}{C - D_j}\frac{\partial e_j}{\partial \beta_D} - \frac{\beta_Y(j)\,\theta_j\, e_j}{(C - D_j)^2}\frac{\partial D_j}{\partial \beta_D}\right] d\beta_D - \left[\frac{\theta_j\, e_j}{C - D_j} + \frac{\beta_Y(j)\,\theta_j}{C - D_j}\frac{\partial e_j}{\partial \beta_Y} + \frac{\beta_Y(j)\,\theta_j\, e_j}{(C - D_j)^2}\frac{\partial D_j}{\partial \beta_Y}\right] d\beta_Y = 0 \tag{A1.3}$$

Totally differentiating **(A1.2)** with respect to β_D and β_Y , we get

$$-\left[\frac{\beta_Y(j)\theta_j}{C-D_j}\frac{\partial D_j}{\partial \beta_D}+M\frac{\partial e_j}{\partial \beta_D}\right]d\beta_D$$
$$-\left[-\theta_j \ln\left(C-D_j\right)+\frac{\beta_Y(j)\theta_j}{C-D_j}\frac{\partial D_j}{\partial \beta_Y}+M\frac{\partial e_j}{\partial \beta_Y}\right]d\beta_Y=0 \qquad \textbf{(A1.4)}$$

The left hand side of **(A1.3)** and **(A1.4)** will be zero for all arbitrary and small values for $d\beta_D$ and $d\beta_Y$. Therefore, the coefficients of $d\beta_D$ and $d\beta_Y$ must always be zero. Equating coefficients of $d\beta_D$ and $d\beta_Y$ in **(A1.3)** and in **(A1.4)** to zero, we get,

$$1-\frac{\beta_Y(j)\theta_j}{C-D_j}\frac{\partial e_j}{\partial \beta_D}-\frac{\beta_Y(j)\theta_j\, e_j}{\left(C-D_j\right)^2}\frac{\partial D_j}{\partial \beta_D}=0 \qquad \textbf{(A1.5)}$$

$$\frac{\theta_j\, e_j}{C-D_j}+\frac{\beta_Y(j)\theta_j}{C-D_j}\frac{\partial e_j}{\partial \beta_Y}+\frac{\beta_Y(j)\theta_j\, e_j}{\left(C-D_j\right)^2}\frac{\partial D_j}{\partial \beta_Y}=0 \qquad \textbf{(A1.6)}$$

$$\frac{\beta_Y(j)\theta_j}{C-D_j}\frac{\partial D_j}{\partial \beta_D}+M\frac{\partial e_j}{\partial \beta_D}=0 \qquad \textbf{(A1.7)}$$

$$-\theta_j \ln\left(C-D_j\right)+\frac{\beta_Y(j)\theta_j}{C-D_j}\frac{\partial D_j}{\partial \beta_Y}+M\frac{\partial e_j}{\partial \beta_Y}=0 \qquad \textbf{(A1.8)}$$

Multiplying both sides of **(A1.7)** with $\frac{e_j}{C-D_j}$ and adding to **(A1.5)**.

$$1-\frac{\beta_Y(j)\,\theta_j}{C-D_j}\frac{\partial e_j}{\partial \beta_D}+M\,\frac{e_j}{C-D_j}\frac{\partial e_j}{\partial \beta_D}=0$$

$$\Rightarrow \qquad \frac{\partial e_j}{\partial \beta_D}=-\frac{C-D_j}{M\,e_j-\beta_Y(j)\,\theta_j} \qquad \textbf{(A1.9)}$$

From **(A1.2)**, we see that the denominator on the right-hand side of **(A1.9)** is positive. The numerator on the right-hand side of **(A1.9)** is also positive by our construction.

$$\therefore \frac{\partial e_j}{\partial \beta_D}<0$$

From **(A1.7)**, we can see that $\frac{\partial D_j}{\partial \beta_D}$ and $\frac{\partial e_j}{\partial \beta_D}$ are of opposite sign.

$$\therefore \frac{\partial D_j}{\partial \beta_D}>0$$

Multiplying both sides of **(A1.8)** with $\frac{e_j}{(C-D_j)}$ and subtracting from **(A1.6)**.

$$\frac{\theta_j\,e_j}{C-D_j}+\frac{\beta_Y(j)\,\theta_j}{C-D_j}\frac{\partial e_j}{\partial \beta_Y}+\theta_j\,\frac{e_j}{C-D_j}\ln(C-D_j)-M\,\frac{e_j}{C-D_j}\frac{\partial e_j}{\partial \beta_Y}=0$$

$$\Rightarrow \frac{\partial e_j}{\partial \beta_Y}=\frac{\theta_j\,e_j}{M\,e_j-\beta_Y(j)\,\theta_j}\left[1+\ln(C-D_j)\right]>0 \qquad \textbf{(A1.10)}$$

From **(A1.6)** we can see that

$$\frac{\partial D_j}{\partial \beta_Y}<0 \qquad \blacksquare$$

Lemma A1.2

The indifference curves in the (β_Y, β_D) plane for both the H-type and L-type agents are concave and downward sloping

Proof:

Let U_j denote the conditional expected utility to an agent of type j, conditional on his/her knowledge of own type j.

$$\therefore U_j = \beta_0(j) + \beta_D(j)D_j + \beta_Y(j)\theta_j e_j \ln(C - D_j) - \tfrac{1}{2}Me_j^2 \qquad \textbf{(A1.11)}$$

$$\begin{aligned}
\therefore \frac{\partial U_j}{\partial \beta_D(j)} &= D_j + \beta_D(j)\frac{\partial D_j}{\partial \beta_D(j)} + \beta_Y(j)\theta_j \ln(C - D_j)\frac{\partial e_j}{\partial \beta_D(j)} \\
&\quad - \frac{\beta_Y(j)\theta_j e_j}{(C - D_j)}\frac{\partial D_j}{\partial \beta_D(j)} - M e_j \frac{\partial e_j}{\partial \beta_D(j)} \\
&= D_j + \left[\beta_D(j) - \frac{\beta_Y(j)\theta_j e_j}{(C - D_j)}\right]\frac{\partial D_j}{\partial \beta_D(j)} \\
&\quad + \left[\beta_Y(j)\theta_j \ln(C - D_j) - M e_j\right]\frac{\partial e_j}{\partial \beta_D(j)} \\
&= D_j
\end{aligned} \qquad \textbf{(A1.12)}$$

$$
\begin{aligned}
\therefore \frac{\partial U_j}{\partial \beta_Y(j)} &= \beta_D(j)\frac{\partial D_j}{\partial \beta_Y(j)} + \theta_j e_j \ln\left(C - D_j\right) \\
&+ \beta_Y(j)\theta_j \ln\left(C - D_j\right)\frac{\partial e_j}{\partial \beta_Y(j)} - \frac{\beta_Y(j)\theta_j e_j}{\left(C - D_j\right)}\frac{\partial D_j}{\partial \beta_Y(j)} \\
&- M e_j \frac{\partial e_j}{\partial \beta_Y(j)} \\
&= \theta_j e_j \ln\left(C - D_j\right) + \left[\beta_D(j) - \frac{\beta_Y(j)\theta_j e_j}{\left(C - D_j\right)}\right]\frac{\partial D_j}{\partial \beta_Y(j)} \\
&+ \left[\beta_Y(j)\theta_j \ln\left(C - D_j\right) - M e_j\right]\frac{\partial e_j}{\partial \beta_Y(j)} \\
&= \theta_j e_j \ln\left(C - D_j\right) \qquad \textbf{(A1.13)}
\end{aligned}
$$

The last equalities in **(A1.12)** and **(A1.13)** are obtained by noting that the expressions within the brackets are equal to zero by virtue of **(A1.1)** and **(A1.2)**. On the indifference curve, in the (β_D, β_Y) plane, dU_j must equal zero.

From **(A1.11)**

$$
\begin{aligned}
dU_j &= \frac{\partial U_j}{\partial \beta_D(j)} d\beta_D(j) + \frac{\partial U_j}{\partial \beta_Y(j)} d\beta_Y(j) &&= 0 \\
\therefore \quad \frac{d\beta_D(j)}{d\beta_Y(j)} &= -\frac{\dfrac{\partial U_j}{\partial \beta_Y(j)}}{\dfrac{\partial U_j}{\partial \beta_D(j)}} = -\frac{\theta_j e_j \ln\left(C - D_j\right)}{D_j} &&< 0 \qquad \textbf{(A1.14)}
\end{aligned}
$$

$$\frac{d^2\beta_D(j)}{d\beta_Y^2(j)} = -\frac{1}{D_j^2}\left[D_j\left\{\theta_j \ln\left(C-D_j\right)\frac{\partial e_j}{\partial\beta_Y(j)} - \frac{\theta_j e_j}{\left(C-D_j\right)}\frac{\partial D_j}{\partial\beta_Y(j)}\right\}\right.$$
$$\left. -\theta_j e_j \ln\left(C-D_j\right)\frac{\partial D_j}{\partial\beta_Y(j)}\right]$$
$$< 0 \qquad \textbf{(A1.15)}$$

The last inequality in **(A1.15)** follows from **Lemma A1.1**.

Therefore the indifference curves in the (β_Y, β_D) plane are concave and downward sloping.

■

Lemma A1.3

Consider a point in the plane (β_Y, β_D). The slope of the indifference curve through this point for the H-type agent is higher in absolute magnitude than the slope of the indifference curve for the L-type agent passing through the same point.

Proof:

Consider the choices of the agent of type j. The agent chooses D_j and e_j. These choices are implicitly given by **(A1.1)** and **(A1.2)**. Eliminating e_j from these two equations we get an expression for D_j in the following implicit form.

$$\beta_D(j) - \frac{\beta_Y(j)\,\theta_j}{C - D_j}\,\frac{\beta_Y(j)\,\theta_j \ln(C - D_j)}{M} = 0$$

$$\Rightarrow \frac{M\,\beta_D(j)}{\beta_Y^2(j)\,\theta_j^2} - \frac{\ln(C - D_j)}{C - D_j} = 0 \qquad \textbf{(A1.16)}$$

Taking partial derivative of **(A1.16)** with respect to θ_j we get

$$\frac{\partial}{\partial\theta_j}\left(\frac{M\,\beta_D(j)}{\beta_Y^2(j)\,\theta_j^2} - \frac{\ln(C - D_j)}{C - D_j}\right) = 0$$

$$\Rightarrow -\frac{2M\,\beta_D(j)}{\beta_Y^2(j)\,\theta_j^3} - \frac{1}{(C - D_j)^2}\left[(C - D_j)\frac{\left(-\frac{\partial D_j}{\partial\theta_j}\right)}{(C - D_j)} - \left(\ln(C - D_j)\right)\left(-\frac{\partial D_j}{\partial\theta_j}\right)\right] = 0$$

$$\Rightarrow \frac{\left[\ln(C - D_j) - 1\right]}{(C - D_j)^2}\left(\frac{\partial D_j}{\partial\theta_j}\right) = -\frac{2M\,\beta_D(j)}{\beta_Y^2(j)\,\theta_j^3}$$

$$\Rightarrow \frac{\partial D_j}{\partial\theta_j} < 0 \text{ because } \ln(C - D_j) > 1 \text{ for an interior value of } D_j \qquad \textbf{(A1.17)}$$

From **(A1.1)** and **(A1.2)** we can obtain the implicit relationship for e_j by eliminating D_j.

$$\beta_Y(j)\,\theta_j \ln\left(\frac{\beta_Y(j)\theta_j\,e_j}{\beta_D(j)}\right) - M\,e_j = 0 \qquad \textbf{(A1.18)}$$

Taking partial derivative of **(A1.18)** with respect to θ_j we get

$$\frac{\partial}{\partial \theta_j}\left[\beta_Y(j)\,\theta_j \ln\left(\frac{\beta_Y(j)\theta_j\, e_j}{\beta_D(j)}\right) - M\, e_j\right] = 0$$

$$\Rightarrow\ \beta_Y(j)\ln\left(\frac{\beta_Y(j)\theta_j\, e_j}{\beta_D(j)}\right) + \beta_Y(j)\,\theta_j\,\frac{1}{\theta_j} + \beta_Y(j)\,\theta_j\,\frac{1}{e_j}\frac{\partial e_j}{\partial \theta_j} - M\,\frac{\partial e_j}{\partial \theta_j} = 0$$

$$\Rightarrow\ \left[M\, e_j - \beta_Y(j)\,\theta_j\right]\frac{\partial e_j}{\partial \theta_j} = \beta_Y(j)\, e_j\left[1 + \ln\left(\frac{\beta_Y(j)\theta_j\, e_j}{\beta_D(j)}\right)\right]$$

$$\Rightarrow\ \frac{\partial e_j}{\partial \theta_j} > 0 \text{ because } \left[M\, e_j - \beta_Y(j)\,\theta_j\right] > 0 \text{ for all interior } e_j \qquad \textbf{(A1.19)}$$

From **(A1.18)** and **(A1.19)** we can see that

$$\theta_L < \theta_H \Rightarrow e_L < e_H \text{ and } D_L > D_H \text{ and } \frac{\ln(C - D_L)}{D_L} < \frac{\ln(C - D_H)}{D_H} \qquad \textbf{(A1.20)}$$

From **(A1.14)**, **(A1.19)**, **(A1.20)** we can see that

$$\text{Magnitude of Slope of Indifference Curve for H} = \frac{\theta_H\, e_H \ln(C - D_H)}{D_H}$$

$$> \text{Magnitude of Slope of Indifference Curve for L} = \frac{\theta_L\, e_L \ln(C - D_L)}{D_L}$$

■

Proposition 2

Assuming interior solutions to the Agent's problem, the H-type agent will declare less dividend than the L-type agent. The H-type agent will also put more effort than the L-type agent.

Proof:

Let D_L^*, D_H^*, e_L^* and e_H^* denote the first best dividend and efforts.

The subscripts, as before, denote the type of the agent. From **(A1.17)** and **(A1.19)**, we can see that

$$D_H^* < D_L^* \text{ and } e_H^* > e_L^*$$

From (IC′), we have for L,

$$\beta_D(L) - \frac{\beta_Y(L)\theta_L e_L}{C - D_L} = 0$$
$$\beta_Y(L)\theta_L \ln(C - D_L) = M e_L$$

$$1 - \frac{\theta_L e_L^*}{C - D_L^*} = 0$$
$$\theta_L \ln(C - D_L^*) = M e_L^*$$

Simplifying we get,

$$\frac{\beta_D(L)}{\beta_Y(L)} = \frac{e_L}{e_L^*}\frac{(C-D_L^*)}{(C-D_L)} \text{ and } \frac{e_L}{e_L^*} = \frac{\beta_Y(L)\ \ln(C-D_L)}{\ln(C-D_L^*)}$$

$$\Rightarrow \quad \frac{\beta_D(L)}{\beta_Y^2(L)} = \frac{\ln(C-D_L)}{\ln(C-D_L^*)}\frac{(C-D_L^*)}{(C-D_L)} > 1 \qquad \because \frac{\beta_D(L)}{\beta_Y(L)} > 1 \text{ and } \beta_Y(L) < 1$$

$$\Rightarrow \quad \frac{\ln(C-D_L)}{(C-D_L)} > \frac{\ln(C-D_L^*)}{(C-D_L^*)}$$

Let $\quad \phi(\mathrm{D}) = \dfrac{\ln(C-D)}{C-D}$

$$\therefore \quad \phi'(D) = \frac{1}{(C-D)^2}\left[(C-D)\frac{(-1)}{(C-D)} - (-1)\ln(C-D)\right]$$

$$= \frac{1}{(C-D)^2}\left[\ln(C-D)-1\right] > 0$$

$$\therefore \frac{\ln(C-D_L)}{(C-D_L)} > \frac{\ln(C-D_L^*)}{(C-D_L^*)} \Rightarrow D_L > D_L^*$$

Also,

$$\frac{e_L}{e_L^*} = \frac{\beta_Y(L)\ \ln(C-D_L)}{\ln(C-D_L^*)} < 1 \quad \because D_L > D_L^* \Rightarrow \frac{\ln(C-D_L)}{\ln(C-D_L^*)} < 1 \text{ and } \beta_Y(L) < 1$$

Therefore

$$D_H = D_H^* < D_L^* < D_L \text{ and } \mathrm{e_H} = e_H^* > e_L^* > e_L$$

■

Lemma A1.4

This lemma is for the case where agents are characterised by the difference in their initial cash C.

Consider a point in the (β_D, β_Y) *plane. The slope of the indifference curve through this point for the H-type agent is lower in absolute magnitude than the slope of the indifference curve for the L-type agent passing through the same point.*

Proof:

From **(A1.14)** we can see that the slope of the indifference curve in this case is given by

$$\frac{d\beta_D(j)}{d\beta_Y(j)} = -\frac{\theta e_j \ln(C_j - D_j)}{D_j} \qquad \textbf{(A1.21)}$$

We also know from **(A1.16)** that the dividend will be given by

$$\frac{M\beta_D(j)}{\beta_Y^2(j)\theta^2} - \frac{\ln(C_j - D_j)}{C_j - D_j} = 0 \qquad \textbf{(A1.22)}$$

Partially differentiating both sides of **(A1.22)** with respect to C_j

$$\frac{1}{(C_j - D_j)^2}\left[(C_j - D_j)\frac{1}{(C_j - D_j)}\left\{1 - \frac{\partial D_j}{\partial C_j}\right\} - \ln(C_j - D_j)\left\{1 - \frac{\partial D_j}{\partial C_j}\right\}\right] = 0$$

$$\Rightarrow \left\{1 - \frac{\partial D_j}{\partial C_j}\right\} = 0 \Rightarrow \frac{\partial D_j}{\partial C_j} = 1 \qquad \textbf{(A1.23)}$$

We also know from **(A1.18)** that the effort will be given by

$$\beta_Y(j)\,\theta \ln\left(\frac{\beta_Y(j)\theta\, e_j}{\beta_D(j)}\right) = M\, e_j$$

$$\Rightarrow \qquad \frac{\partial e_j}{\partial C_j} = 0 \qquad \textbf{(A1.24)}$$

Therefore from **(A1.21)**, we get

$$\frac{\partial}{\partial C_j}\left(\frac{d\beta_D(j)}{d\beta_Y(j)}\right) = -\frac{\theta\, e_j}{D_j^2}\left[\frac{D_j}{(C_j - D_j)}\left\{1 - \frac{\partial D_j}{\partial C_j}\right\} - \ln(C_j - D_j)\frac{\partial D_j}{\partial C_j}\right]$$

$$= \frac{\theta\, e_j}{D_j^2}\ln(C_j - D_j)$$

■

Lemma A1.5

This lemma is for the case where agents are characterized by the difference in their initial cash C.

For large investment

$$\left|\frac{\partial D_j}{\partial \beta_D}\right| > \left|\frac{\partial D_j}{\partial \beta_Y}\right|$$

Proof:

From **Lemma A1.1** equation **(A1.9)** , we see that

$$\frac{\partial e_j}{\partial \beta_D} = -\frac{C_j - D_j}{M\, e_j - \beta_Y(j)\,\theta} \qquad \textbf{(A1.25)}$$

From **(A1.7)**,

$$\frac{\beta_Y(j)\theta}{C_j - D_j}\frac{\partial D_j}{\partial \beta_D} + M\frac{\partial e_j}{\partial \beta_D} = 0$$

$$\Rightarrow \quad \frac{\partial D_j}{\partial \beta_D} = \frac{C_j - D_j}{\beta_Y(j)\theta} M \frac{C_j - D_j}{M e_j - \beta_Y(j)\theta}$$

$$= \frac{M\left(C_j - D_j\right)^2}{\beta_Y(j)\theta\left(M e_j - \beta_Y(j)\theta\right)} \qquad \textbf{(A1.26)}$$

From **(A1.10)**, we see that

$$\frac{\partial e_j}{\partial \beta_Y} = \frac{\theta e_j}{M e_j - \beta_Y(j)\theta}\left[1 + \ln\left(C_j - D_j\right)\right] \qquad \textbf{(A1.27)}$$

From **(A1.6)**, we see that

$$\frac{\theta e_j}{C_j - D_j} + \frac{\beta_Y(j)\theta}{C_j - D_j}\frac{\partial e_j}{\partial \beta_Y} + \frac{\beta_Y(j)\theta e_j}{\left(C_j - D_j\right)^2}\frac{\partial D_j}{\partial \beta_Y} = 0$$

$$\Rightarrow \quad \frac{\partial D_j}{\partial \beta_Y} = -\frac{C_j - D_j}{\beta_Y(j)\, e_j}\left[e_j + \beta_Y(j)\frac{\partial e_j}{\partial \beta_Y}\right]$$

$$= -\frac{C_j - D_j}{\beta_Y(j)\, e_j}\left[e_j + \beta_Y(j)\frac{\theta e_j}{M e_j - \beta_Y(j)\theta}\left[1 + \ln\left(C_j - D_j\right)\right]\right]$$

$$= -\frac{\left(C_j - D_j\right)\left[M e_j + \beta_Y(j)\theta \ln\left(C_j - D_j\right)\right]}{\beta_Y(j)\left(M e_j - \beta_Y(j)\theta\right)} \qquad \textbf{(A1.28)}$$

From **(A1.26)** and **(A1.28)**, we see that,

$$\frac{\left|\frac{\partial D_j}{\partial \beta_D}\right|}{\left|\frac{\partial D_j}{\partial \beta_Y}\right|} = \frac{M\left(C_j - D_j\right)^2}{\beta_Y(j)\theta\left(M e_j - \beta_Y(j)\theta\right)} \frac{\beta_Y(j)\left(M e_j - \beta_Y(j)\theta\right)}{\left(C_j - D_j\right)\left[M e_j + \beta_Y(j)\theta \ln\left(C_j - D_j\right)\right]}$$

$$= \frac{M\left(C_j - D_j\right)}{\theta\left[M e_j + \beta_Y(j)\theta \ln\left(C_j - D_j\right)\right]}$$

$$= \frac{M\left(C_j - D_j\right)}{2\beta_Y(j)\theta^2 \ln\left(C_j - D_j\right)} \qquad \because M e_j = \beta_Y(j)\theta \ln\left(C_j - D_j\right)$$

$$> 1 \quad \text{for large investment i.e. for large} \left(C_j - D_j\right)$$

■

Proposition 4

This proposition is for the case in which agents are characterized by the difference in their initial cash C. Assuming interior solutions to the agent's problem, if the investment level is high then the H-type agent will declare a larger dividend than the L-type agent.

Proof:

Suppose L is offered a contract $\{\beta_0(L), \beta_D(L), \beta_Y(L)\}$. We know that H would be offered a contract $\{\beta_0(H), \beta_D(H)=1, \beta_Y(H)=1\}$ such that H is indifferent between the contracts $\{\beta_0(L), \beta_D(L), \beta_Y(L)\}$ and $\{\beta_0(H), \beta_D(H)=1, \beta_Y(H)=1\}$. We also know that $\beta_D(L) < \beta_Y(L)$.

Let $D_{H|L}$ be the dividend level that would have been selected by H if he/she was offered the contract presented to L. We know from **(A1.23)** that $D_{H|L} > D_L$.

Suppose we offer a contract $\{\beta_0(L)+\Delta\beta_0,\ \beta_D(L)+\Delta\beta_D,\ \beta_Y(L)+\Delta\beta_Y\}$ to H, such that H is indifferent between this contract and the contract offered to L, $\Delta\beta$ s are small and $\Delta\beta_D > \Delta\beta_Y$. Let the new dividend level chosen by H be $D_{H|L}+\Delta D$.

$$\therefore \quad \Delta D = \frac{\partial D}{\partial \beta_D}\Delta\beta_D + \frac{\partial D}{\partial \beta_Y}\Delta\beta_Y > 0 \qquad \textbf{(A3.29)}$$

From **Lemma A1.1** and **Lemma A1.5** , we see that $\Delta D > 0$.

Consider the indifference surface of H in the $\{\beta_0,\ \beta_D,\ \beta_Y\}$ space passing through the points $\{\beta_0(L),\ \beta_D(L),\ \beta_Y(L)\}$ and $\{\beta_0(H),\ \beta_D(H)=1,\ \beta_Y(H)=1\}$. Consider a continuous path on this indifference surface joining the two points $\{\beta_0(L),\ \beta_D(L),\ \beta_Y(L)\}$ and $\{\beta_0(H),\ \beta_D(H)=1,\ \beta_Y(H)=1\}$ such that as we move from $\{\beta_0(L),\ \beta_D(L),\ \beta_Y(L)\}$ towards $\{\beta_0(H),\ \beta_D(H)=1,\ \beta_Y(H)=1\}$ along this path, the change in β_D is more than the change in β_Y. Such a path exists because $\beta_D(L) < \beta_Y(L)$. By virtue of **(A3.29)** we see that $D_H > D_{H|L}$. Therefore $D_H > D_L$.

■

Appendix II: A Simple Model

In this appendix, we examine a much simpler model where we leave out effort from our production function. Specifically, the production function considered here is $\tilde{Y}_j = \theta_j \ln(C-D_j) + \tilde{\varepsilon}$, where $\tilde{\varepsilon}$ is random noise with zero mean. The other notation is the same as that used in Chapter 3. For the sake of easy reference, we reproduce the notation here again.

C Cash available for investment

I_j Investment level chosen by agent of type j, when agent truthfully declares type to be j.

D_j Dividend level Selected by an Agent of type j, when agent truthfully declares type to be j = $C-I_j$.

θ_j Productivity parameter for agent of type j. For our purpose, we assume that j can only have two values - L(ow) and H(igh) and $\theta_L < \theta_H$.

$\tilde{\omega}_{\hat{j}}$ The wage function offered to an agent, assumed to be linear in D and $\tilde{Y}$. Thus $\tilde{\omega}_{\hat{j}} = \beta_0\left(\hat{j}\right) + \beta_D\left(\hat{j}\right)D_j + \beta_Y\left(\hat{j}\right)\tilde{Y}$

The analysis can be done in the same way as it was done in Chapter 3. We first note that the first-best solution would mean solving the following problem:

$$\underset{D_j}{Maximize} \qquad D_j + \theta_j \ln\left(C - D_j\right) \qquad \forall j$$

The first order condition would be

$$1 - \frac{\theta_j}{C - D_j^*} = 0$$

$$\Rightarrow \qquad D_j^* = C - \theta_j \qquad \textbf{(A2.1)}$$

where D_j^* is the first best level of dividends.

In the second best solution, the H-type would be offered a contract which sets $\beta_D = \beta_Y = 1$; that is, a contract which "sells" the company to the H-type. The dividend level selected by H-type would be the first-best dividend level. The contract that would be offered to L-type would be an inefficient contract. The inefficient contract would lead L to depart from the first-best optimum, but would limit the information rent paid to the H-type.

In order to determine the nature of this inefficient contract, we need to examine the nature of the indifference curve for L and H in the (β_Y, β_D) plane. We note that for a given value of (β_Y, β_D), an agent solves the following problem.

$$\underset{D_j}{Maximize} \quad \beta_D D_j + \beta_Y \theta_j \ln\left(C - D_j\right) \qquad \text{for every j}$$

Let $U_j \equiv \beta_D D_j + \beta_Y \theta_j \ln\left(C - D_j\right)$

First order condition is

$$\beta_D - \frac{\beta_Y \theta_j}{C - D_j} = 0$$

$$\Rightarrow \qquad D_j = C - \frac{\beta_Y}{\beta_D}\theta_j \qquad \textbf{(A2.2)}$$

Also,

$$\frac{\partial U_j}{\partial \beta_D} = D_j \text{ and } \frac{\partial U_j}{\partial \beta_Y} = \theta_j \ln\left(C - D_j\right)$$

$\Rightarrow$ Slope of the indifference curve in (β_Y, β_D) plane

$$= \quad -\frac{\dfrac{\partial U_j}{\partial \beta_Y}}{\dfrac{\partial U_j}{\partial \beta_D}} = -\frac{\theta_j \ln\left(C - D_j\right)}{D_j} \qquad \textbf{(A2.3)}$$

By taking the total differential of the first order condition with respect to β_D and β_Y , we get

$$\left[1 - \frac{\beta_Y \theta_j}{\left(C - D_j\right)^2}\frac{\partial D_j}{\partial \beta_D}\right] d\beta_D - \frac{\theta_j}{C - D_j}\left[1 + \frac{\beta_Y}{\left(C - D_j\right)}\frac{\partial D_j}{\partial \beta_Y}\right] d\beta_Y = 0$$

The left hand side of the above equation will be zero for all arbitrary and small values for $d\beta_D$ and $d\beta_Y$. Therefore, the coefficients of $d\beta_D$ and $d\beta_Y$ must always be zero. Equating coefficients of $d\beta_D$ and $d\beta_Y$ to zero and simplifying, we get

$$\frac{\partial D_j}{\partial \beta_D} = \frac{\left(C - D_j\right)^2}{\beta_Y \theta_j} > 0, \text{and}$$

$$\frac{\partial D_j}{\partial \beta_Y} = -\frac{C - D_j}{\beta_Y} < 0$$

Differentiating the slope of the indifference curve **(A2.3)** with respect to β_Y and using the relation above, we see that in the (β_Y, β_D) plane the indifference curve for an agent is downward sloping and concave. By partially differentiating the first order condition **(A2.2)** with respect to θ_j , we get

$$\frac{\partial D_j}{\partial \theta_j} = -\frac{\beta_Y}{\beta_D} < 0$$

Differentiating **(A2.3)** with partially with respect to θ_j, we see

$$\frac{\partial}{\partial \theta_j}(\text{Slope of the indifference curve})$$

$$= -\frac{1}{D_j^2}\left[D_j\left\{\ln(C-D_j)+\frac{1}{C-D_j}\left(-\frac{\partial D_j}{\partial \theta_j}\right)\right\}-\left\{\theta_j \ln(C-D_j)\right\}\frac{\partial D_j}{\partial \theta_j}\right]$$

$$< 0 \qquad \because \frac{\partial D_j}{\partial \theta_j} < 0$$

From above we know that the sorting condition is satisfied and that at any point in the (β_Y, β_D) plane the indifference curve for H type is steeper than the indifference curve for L type passing through the same point. This means that the indifference surface will be like the one depicted in Figure 1 and Figure 2. Using arguments similar to the ones used in Chapter 3, we can see that

$$\frac{\beta_D(L)}{\beta_Y(L)} > 1, \beta_D(L) < 1 \text{ and } \beta_Y(L) < 1$$

$$\beta_D(H) = 1 \text{ and } \beta_Y(H) = 1$$

Now, $\theta_H > \theta_L$

$$\Rightarrow \quad \theta_H > \frac{\beta_Y(L)}{\beta_D(L)}\theta_L$$

$$\Rightarrow \quad C - \theta_H < C - \frac{\beta_Y(L)}{\beta_D(L)}\theta_L$$

$$\Rightarrow \quad D_H < D_L \qquad \text{from } \textbf{(A2.2)}$$