ESSAYS IN CORPORATE FINANCE

By

JEFFREY CHARLES COLPITTS

B.B.A., Bishop's University, 1999

M.Sc. Business, The University of British Columbia, 2002

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In

THE FACULTY OF GRADUATE STUDIES

BUSINESS ADMINISTRATION

THE UNIVERSITY OF BRITISH COLUMBIA

March 2007

Abstract

In the first essay, I consider the impact of tort liability on firms capital structure. Tort litigation is not only a substantial risk facing firms worldwide, but is also a unique form of risk, in that it can be exacerbated or mitigated by how firms adjust their debt-equity mix. I examine how firms ought to adjust their capital structure when faced with litigation, and consider various extensions to basic model. These include the interaction between capital structure, tort liability and insurance, how the problem changes when several firms face tort risk and are jointly and severally liable, and the implications that arise from moving from a one period to a two period setting.

In the second essay, we develop and test a theory of insurers' choice of the mix of equity and liabilities. The role of equity in insurance markets and in our model is to back insurers' promises to pay claims when there is aggregate uncertainty, or dependence among risks. Depending on the nature of this aggregate uncertainty, the equity held by firms in a competitive insurance market may increase with rising uncertainty, or it may initially increase then decrease. The ratio of equity to revenue unambiguously increases with uncertainty. We test the model, as well as implications of recent models of insurance market dynamics, on a cross-section of U.S. property-liability insurers.

In the third essay, I examine optimal contracting with risk averse managers. I start from the following observations: (1) managers select projects and exert effort; (2) risk averse managers make distorted project selection decisions, and this problem is increasing in risk aversion; (3) managers with low risk aversion are attracted to high-power compensation packages. I develop a model where high-power incentive contracts act as screening devices, helping firms attract less risk averse managers who will then make less distorted project selection decisions. Optimal contracts trade off the screening and effort-inducing

benefits of incentive contracts against the deviation from optimal risk sharing. The resulting equilibrium provides a new perspective on why some managerial contracts feature such high-powered incentives, as well predictions for the cross-sectional variation in the power of incentive contracts.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank the faculty members of the Sauder School of Business who helped me throughout the PhD program. In particular, I would like to thank Harjoat Bhamra, Gilles Chemla, Ruth Freedman, Ron Giammarino, Marcin Kacperczyk, Mo Levi, Alan Kraus, Cornelia Kullmann, Kai Li, and Tan Wang for helpful advice throughout my studies – some related to these essays, and some not. I would particularly like to thank Murray Carlson and Rob Heinkel for taking the time to serve on my thesis committee.

I would also like to thank classmates past and present whose friendship and great ideas were a source of inspiration: Chris Bradley, Casey Clements, Julian Douglass, Shinsuke Kamoto, Caglar Kamu, Lars Kuehn, Kenny Pun, Thomas Ruf, Jan Schneider, Daniel Smith, Issouf Soumaré, Bill Stewart, Richard Taylor, Marcus Xu and Longkai Zhao.

I would like to make special mention of my thesis advisor, Ralph Winter. There are no words sufficient to describe the positive impact his influence has had on me, academically, professionally and personally, so I won't grasp for them. It is not an exaggeration to say that this thesis would not have been completed had it not been for his well-timed intervention.

In the end it is impossible to name people who helped me complete this thesis without overlooking many more whose names could just as well be on this page. To those whom I have missed, I apologize.

Finally, I would like to dedicate this thesis to my family, whose interest in corporate finance is negligible, but whose interest in me finishing this work has been great, and much appreciated.

Statement of Co-Authorship

The second essay (Chapter 3) is written jointly with Ralph Winter and Dajiang Guo. My contributions to the paper included building and refining the theoretical framework, gathering the data, analyzing the data, and preparing the final manuscript.

Chapter 1

Introduction

In this thesis I examine three different topics in corporate finance. The exact nature of the questions posed differ across the three essays, but they are all questions of importance to business entities.

In the first essay, I examine how firms faced with tort liability ought to adjust their capital structure, the mix between debt and equity securities. Tort liability has expanded enormously over the years. Firms must be aware that a substantial portion of the firm's assets and cash flows are at risk of being transferred to tort claimants, should they win a legal judgment against the firm. The law and economics literature has advanced a good deal of study to ways in which firms ought to seek to mitigate this problem. One possibility that has received scant notice from scholars working in the area is that firms are able to alter their potential exposure to tort risk by making changes to their capital structure. Specifically, corporations facing potential legal risk may do better to finance themselves with a greater proportion of debt than they would otherwise.

Most operational risks affect the total value of the firm's cash flows regardless of how claims to those cash flows are structured. This is not the case with tort risk. When tort claimants win a judgment against a firm, the amount they recover depends on what other promises have been made regarding how the firm's assets will be distributed. In terms of priority of claims, in most jurisdictions, tort claimants collect ahead of shareholders, but have equal or lower priority than unsecured creditors, and lower priority than secured creditors. This implies that when tort claimants win a judgment against a firm, they can collect fully against equity holders. However, once the equity share has been exhausted and the firm forced into bankruptcy, the tort claimants either share on a pro rata basis with unsecured creditors after secured creditors have been paid fully, or tort claimants make no further collection whatsoever.

I develop a simple model where a firm trades off the asset-shielding and tax advantages of debt against the increased probability of bankruptcy costs. I then consider extensions to the model, the first considering the availability of liability insurance, the second considering joint and several liability regimes.

In the second essay, we explore the cross-sectional variation in insurers' capital structures: the choice by stock insurers of the mix of equity and liability. As in the standard theory of optimal capital structure in finance, predictions of the theory must rely on specific capital market imperfections. We focus here on the simplest one: that issuing and maintaining additional equity is costly. Our model yields testable implications with a focus (appropriate for an analysis of insurance markets) on the liability side of the market.

We develop the simplest model of an insurance market with costly equity, in a two-period setting. For equity to have any role in an insurance market there must be aggregate uncertainty, or dependence among insured risks; the absence of a law of large numbers means that equity is necessary to back up promises to pay claims in the event of adverse realizations of aggregate shocks. Accordingly, the key comparative static issue that we focus on is the impact of increasing aggregate uncertainty.

We test the theory using cross-sectional data on U.S. property-liability insurers. The focus is on tests of two hypotheses. The first is the implication of the static model, that leverage is decreasing in aggregate uncertainty. The second is an implication of previous dynamic models of competitive insurance markets that external equity is more costly than internal equity – specifically that there is a positive cost to the "round-trip" of distributing an amount of cash then raising the same amount in external equity. We also offer a link between the recent insurance market literature and corresponding empirical results in tests of capital structure for non-financial corporations.

In the third essay, I consider the problem firms face when contracting with managers when there is heterogeneity in risk aversion in the pool of managerial labour. I motivate the essay with a number of observations. The first is that managers differ in their degree of risk aversion, and that a manager's risk aversion is not observable.

Second, higher managerial risk aversion is costly in two ways. First, higher risk aversion means that the manager puts a lower value on risky pay. This implies that the cost of motivating effort exertion is increasing in managerial risk aversion. However, managerial risk aversion is also costly in terms of motivating correct *project selection.* When selecting projects, managers have an incentive to choose those that best fit their own interests, as opposed to those of firm shareholders. The greater difference in risk preferences between risk averse managers and risk neutral shareholders, the greater will be the distortion imposed by managers selecting projects according to their own interests.

A third observation is that the market for managerial labour, like any labour market, is a competitive one. Firms compete with one another for the services of preferred managers, and managers will choose to work for the firm that makes them the offer they prefer.

These three observations taken together imply the following. From the second observation, it is clear that firms prefer lower risk aversion managers. From the third observation, they must compete against other firms for the services of lower risk aversion managers. And from the first observation, such competition is difficult, since a manager's risk aversion is his own private information. Firms must therefore develop contracts which serve as screening devices, designed so that they will attract low risk aversion managers. Since all managers prefer more pay to less, firms cannot compete for low risk aversion managers simply by raising wages. Since lower risk aversion managers put greater value on risky pay than high risk aversion managers, firms have an incentive to offer high-powered contracts as a screening device. Such contracts have greater appeal to the targeted low risk aversion managers.

I develop a model where firms must compete against one another in the managerial labour market to attract managers who are responsible for both project selection and effort exertion. In this setting, incentive contracts perform two functions. The first is to serve the traditional role of motivating the correct effort choice. The second is to act as a screening mechanism, helping firms compete for the services of a lower risk aversion manager whose preferences lead to better project selection.

Chapter 2

Tort Liability and Capital Structure

2.1 Introduction

Firms face risk from a variety of sources. One type of risk that has expanded enormously over the years is that posed by legal liability. Operating in an increasingly litigious society means that firms must be aware that a substantial portion of the firm's assets and cash flows are at risk of being transferred to tort claimants, should they win a legal judgment against the firm. The law and economics literature has advanced a good deal of study to ways in which firms ought to seek to mitigate this problem. The ideas range from working to avoid lawsuits in the first place by exercising greater care, to purchasing insurance in order to substitute a sure loss for exposure to the stochastic whims of juries, to restructuring the firm so that there are fewer assets exposed to legal liability. One possibility that has received scant notice from scholars working in the area is that firms are able to alter their potential exposure to tort risk by making changes to their capital structure. Specifically, corporations facing potential legal risk may do better to finance themselves with a greater proportion of debt than they would otherwise.

Since Modigliani and Miller first posited that under a set of restrictive assumptions a firm's capital structure does not matter, numerous models have emerged which attempt to demonstrate alternative circumstances under which a firm's capital structure might indeed impact the aggregate value of securities issued by a firm. To the extent that firms have an optimal or "target" capital structure, it is most commonly modeled as a tradeoff between some tax advantage provided by debt, versus some increased probability that the firm will be bankrupt and incur bankruptcy costs. This is the standard tradeoff model.

One way of discussing the standard tradeoff model is to consider the various parties' claims to firm cash flows. The firm's goal, when choosing its capital structure, is to maximize the value of claims belonging to various groups of security holders (usually, bondholders and shareholders, although more complex forms are possible). Maximizing security holders' claims to assets entails minimizing the value of claims that will accrue to other parties, such as the government (taxes) and direct or indirect bankruptcy costs.

The tradeoff model has had little to say about the specific type of risk posed by a tort judgment. Most operational risks affect the total value of the firm's cash flows regardless of how claims to those cash flows are structured. This is not the case with tort risk. When tort claimants win a judgment against a firm, the amount they recover depends on what other promises have been made regarding how the firm's assets will be distributed. In terms of priority of claims, in most jurisdictions, tort claimants collect ahead of shareholders, but have equal or lower priority than unsecured creditors, and lower priority than secured creditors. This implies that when tort claimants win a judgment against a firm, they can collect fully against equity holders. However, once the equity share has been exhausted and the firm forced into bankruptcy, the tort claimants either share on a pro rata basis with unsecured creditors after secured creditors have been paid fully, or tort claimants make no further collection whatsoever.

Thus the effect of tort risk on capital structure involves a tradeoff of its own. To the extent that tort liability adds risk to cash flows, and decreases the expected value of cash flows that can be paid to other parties, an increase in tort liability increases the risk of bankruptcy for a given debt level. This may induce firms to reduce debt. On the other hand, as firms' increase the level of debt in their capital structure, the more likely that the firm can take advantage of tort claimants' relatively low priority to reduce the amount that they are expected to be paid. This countervailing effect sees firms increasing debt as tort liability increases.

The purpose of this paper is to determine how firms best ought to use debt, in the face of tort liability, to maximize the aggregate value of the firm's securities. I develop a simple model where a firm trades off the asset-shielding and tax advantages of debt against the increased probability of bankruptcy costs. I then consider extensions to the model. In the first, the firm has the option to purchase liability insurance,in addition to setting a debt level, to mitigate total expected judgment, tax and bankruptcy costs. In the second extension, I explore a situation where two firms face judgment risk in a joint and several liability regime. In this setting, each firm is liable for one half of the total judgment, plus whatever portion the other firm is unable to pay. The section considers how the interaction between the two firms' capital structure decisions affects the equilibrium debt level. Finally, using a two-period version of the model, I consider how firms' optimal target debt level evolves as tort risk changes over time.

2.2 Motivation and Literature Review

2.2.1 Why is tort liability important, and to which firms?

It is generally accepted that the number of lawsuits in the United States has seen a remarkable increase over the years. Tillinghast Towers Perrin estimates that 2004

U.S. tort costs exceeded $260 billion, or 2.22% of US GDP. That figure represents an average per capita cost of $886 for every citizen of the United States. Since 1950, the average annual percent increase in total tort costs has exceeded annual GDP growth by more than 2 percent. Commercial tort costs, the type most relevant to this paper, have grown at the fastest rate of late. From 1999 to 2004, commercial tort costs increased at an average annual rate of 11.6% per year. (Tillinghast Towers Perrin, 2006)

An extremely litigious society, coupled with juries that over the years have awarded hefty punitive damages with increasing enthusiasm, means that all economic actors are aware of the substantial risk posed by the potential of tort litigation.

Not all firms are equally likely to face a lawsuit. Certain lines of business naturally engender greater risk of imposing harm on others, and suffering a judgment as a result. Tobacco, waste management, firearms, chemical manufacturing, medical devices and pharmaceuticals are examples of the industries where the very nature of the business leads to risk of imposing harm on others. This in turn leads to the potential for litigation on a massive scale. States' 1990's litigation against the large tobacco companies, and a slew of recent class action suits against Merck for the manufacture of Vioxx are just two examples.

Firms affected by asbestos litigation have been hit particularly hard. Over 6000 firms have faced asbestos-related lawsuits, and the vast majority of these firms were not involved in the manufacture of asbestos products. In more than 60 cases, the litigation led directly to the defendant firm's bankruptcy. (Carroll et al 2002)

While tort risk varies across industries, no firm is immune, and operating in any line of business can lead to litigation. The Loewen Group, an aggregator of funeral homes, was involved in what appeared to be a minor dispute over a few million dollars in service contracts. This situation eventually led to a $500M judgment against Loewen, bankrupting the firm, by the time a Mississippi jury was done deliberating the case. While firms in certain industries are at particular risk of finding themselves defending tort claims in court, the fate can potentially befall almost any firm.

2.2.2 Why tort liability matters from a capital structure perspective

An argument that tort liability deserves special consideration when considering a firm's capital structure must include an explanation of how tort liability differs from other forms of risk the firm faces. If tort liability were merely a stochastic reduction in the firm's terminal cash flows, analyzing it separately from other forms of cash flow risk would not yield any particular insight.

The important distinction with respect to tort liability risk is that its potential impact on firm value depends on how claims to the firm's cash flows are structured. For example, unlike a $100M reduction in the value of firm assets due to changing product market conditions, a \$100M judgment against a firm does not necessarily reduce the value of a firm's assets by \$100M. Tort claimants can collect their judgment only up to the value of equity securities. Once the firm's equity is exhausted and the firm is forced into bankruptcy, tort claimants are left to collect as much as their claim as they can from the firm's assets after more senior creditors have been paid.[1]

In the standard tradeoff model, a firm faced with a negative stochastic impact to cash flows will tend to move to a lower debt level. Expected tax savings are diminished, and the probability of facing bankruptcy is increased. This is not necessarily the case with tort liability. The dollar value of firm assets actually paid to tort claimants is limited to the value of the assets not promised to higher priority claimants - for the purposes of this paper, the debt holders. Thus while increased tort liability brings with it the increased probability of bankruptcy, and therefore the increased probability of incurring bankruptcy costs, it also brings greater potential savings due debt. In addition to debt providing a tax advantage, for a firm faced with tort liability it also provides an "asset shielding" benefit: a dollar of cash flows promised to debt holders cannot be fully expropriated to pay tort claimants.[2]

[1]See Painter (1984) for a detailed description of tort claimant priority with respect to other creditors. Depending on the jurisdiction and other circumstances, tort claimants have either (a) lower priority than secured creditors and equal priority to unsecured creditors, or (b) lower priority than all debtholders. For the purposes of this paper, tort creditors are assumed to have lower priority than debtholders. "Debt" as described in this paper should therefore be interpreted as being an instrument that gives its holder higher priority than tort claimants.

[2]The comparison between the standard tradeoff model and a model involving tort risk starts in Section 3 with the simplest comparative static results.

The fundamental difference between an operational risk that may potentially negatively impact cash flows and tort liability is that - unlike tort risk- the potential loss from the operational risk cannot be mitigated by adjusting the firm's relative amounts of debt and equity financing. While optimal capital structure is affected by the operational risk, the operational risk to cash flows is not affected by capital structure. On the other hand, the magnitude of potential tort liability *is* affected by capital structure. When considering how to deal with tort risk, the firm doesn't only consider the costs that would come with financial distress; financial distress provides the advantage of shielding some of the firm's cash flows from tort claimants' reach, reducing the ex ante value of their claim. Because these two effects work in opposite directions, the direction of the impact of tort risk on optimal capital structure is not immediately obvious.

In other ways expected tort liability is similar to a firm's expected tax liability. Both depend on the firm's capital structure, and represent expected claims on cash flows to be paid to parties other than a firm's security holders. However, a firm's tax liability is not in itself stochastic. While the exact realization of a firm's tax bill is uncertain ex ante, it is a deterministic function of the firm's eventual cash flows, promised payments to debt holders, and the residual cash flows accruing to equity holders. A firm's expected payment to tort claimants also depends on the firm's cash flow and the relative mix of debt and equity. Tort liability is not a deterministic function, in that the future decisions of judges and juries are uncertain. Another crucial difference is that tax liability does not tend to push a firm into bankruptcy, while an unfavourable tort judgment most certainly can.[3]

Finally, as I demonstrate in Section 5, tort risk involves interesting interactions or externalities among firms in their capital structure decisions. This moves capital structure from the realm of a single agent decision to game theory. The externalities give rise to multiple equilibria, where aggregate debt levels can end up being much higher or much lower than they would be in the absence of the interaction between agents' decisions.

2.2.3 Tort liability and bankruptcy

To date, there is little in the way of research into how firms adjust capital structure when faced with tort liability. However, there are numerous cases, many high profile, where tort judgments have pushed firms into bankruptcy. This possibility must be taken into account when firms determine their capital structure. Any assumption to the contrary strains credulity.

Among the most high profile instances of a firm going bankrupt as a result of lawsuits is the Johns-Manville company. One of the earliest cases of a "mass tort", the asbestos manufacturer soon became deluged by lawsuits from plaintiffs alleging health problems as a result of exposure to the firm's product. In 1982, Manville filed

[3]Of course, it is possible that an unpaid tax bill on a firm's past earnings could lead to the firm being forced into bankruptcy. However, this situation would be analogous to the tax collector winning a "judgment" against the firm, and would therefore fit into the model as a form of tort risk liability.

for bankruptcy under Chapter 11, "not because of an inability to meet its current debts, but rather because of its anticipation of massive asbestos personal injury claim liability in the future" (Vairo 2003)

While Manville was one of the first firms to go bankrupt as a result of asbestos liability claims, it was certainly not the last. According to the Rand Institute for Civil Justice, over 6000 firms in nearly every industry have faced at least one lawsuit related to asbestos liability. For most firms the cost is negligible relative to firm assets, but over 60 firms have filed for bankruptcy as a direct result of asbestos litigation. (Carroll et al 2002)

And cases where litigation leads to bankruptcy are not limited to asbestos. Other high profile cases include that of Dow Corning, which filed for bankruptcy in 1992, awash in litigation stemming from injuries caused by breast implants, and A.H. Robins, which filed for bankruptcy in 1985 as a result of litigation related to its Dalkon Shield intrauterine device. There is little in the way of comprehensive evidence linking tort liability and bankruptcy. However, a wealth of anecdotal evidence suggests that this type of problem does occur frequently, and to large firms, and thus motivates an examination of the logic of optimal capital structure for firms facing tort risk.

2.2.4 Judgment proofing

Using capital structure to reduce exposure to tort liability risk is only one means by which firms are able to reduce their exposure to lawsuits. Other methods exist, with one of the most common, and most commonly studied, is to create "narrow entities". That is, to the extent that certain risky lines of business likely to lead to tort liability can be isolated from the rest of the firm, then this is what the firm should do.[4] If the risky behaviour gives rise to a lawsuit, then plaintiffs will be left to sue an entity whose pockets are much less deep than those of the firm as originally constituted.

Ringleb and Wiggins find evidence related to this form of judgment proofing. Their proxy for lawsuit liability is industry worker exposure to carcinogens. After controlling for various other possible explanatory factors, they find that the proportion of small firms operating in a given industry is significantly positively correlated with the degree of worker carcinogen exposure in that industry. They take this as evidence that firms operating in industries associated with potential judgment risk tend to be smaller, and therefore have fewer assets.

The propensity of firms in risky industries to operate as narrowly as possible does not obviate the need to consider asset shielding through capital structure as an alternative, or in some cases complementary, technique. In some instances, it may not be possible to separate risky activities from less risky activities. While it was

[4] See Lopucki (1998) and Roe (1986) for discussion of these ideas in detail.

logical for the RJ Reynolds tobacco business to be split from the Nabisco division, it would be impossible for RJ Reynolds to further separate sales of the cigarettes that cause cancer from those that don't. When a risky activity is the core of the firm, further separation is simply not feasible.

Further, in some instances, a cost must be incurred to set up narrow entities. If a risky activity is integral to a firm's broader activities, and the efficiencies from keeping the risky division internally outweigh the foregone expected tort judgment savings from not spinning it out, the firm will keep the division internally.

There may be legal impediments to judgment proofing. For example, legislators may mandate that firms performing certain activities have sufficient resources to pay potential litigants in the event that they cause a tort. This requirement would normally be satisfied either by minimum asset requirements or by compulsory liaibility insurance.[5]

As well, U.S. law provides a means of reducing the advantage of judgment proofing through the creation of a narrow entity. Courts have the power to "pierce the corporate veil"; that is, in some instances courts hold shareholders of a tortfeasor firm liable beyond the value of their shares. This is most likely to occur exactly when the firm has structured itself narrowly. According to Bergmann, "corporations are expected to operate with a certain minimal level of assets that takes into account

[5]See Shavell (2005) for a more detailed discussion of these types of solutions to judgment-proofing problems, as well as an analysis of how such requirements affect incentives with respect to making care decisions to avoid accidents in the first place.

the particular nature and risks of that enterprise". (Bergmann 2004) In other words, creating a narrow entity solely for the purpose of performing risky tasks in order to shield assets against simply may not work.

Finally, even in cases where as narrow an entity as possible is established, and assuming that the entity has been established in such a way that the courts will not engage in veil piercing, that entity will still face potential tort judgment liability. That firm must make a capital structure decision in the presence of that liability, making the research questions posed by this paper relevant for that firm.

2.2.5 Previous research

In the law and economics literature, there has been some research that considers the role bankruptcy (and by extension, capital structure) has to play in the context of a firm that faces tort liability. However, the focus has largely been on how tortfeasors behave given the potential for insolvency. Huberman et al (1983) consider how an economic entity will make an insurance decision when liability has the potential to make it insolvent. They find that bankruptcy protection leads firms to lower levels of insurance than would be optimal otherwise.[6] Kornhauser and Revesz (1990) consider how the potential for insolvency will affect a firm's decision as to the level of care it will take to avoid incurring a lawsuit, under different liability regimes. These, and

[6]The result is driven by the notion that an insurer must pay for damages caused by the firm even in states where the firm is insolvent. As such, when insurance is fairly priced from the insurer's perspective, the firm pays for coverage even in bankruptcy states where it has no need for insurance. Because the firm is paying "too much", it has an incentive to back away from full insurance.

other similar papers, generally takes the probability that the firm becomes insolvent to be exogenous. That is, the firm does not make a capital structure decision in these models. As I make clear in this paper, this is problematic, as capital structure is an endogenous decision made in the context of all risks facing the firm, including tort risk.

Other papers consider how different regulations regarding lender liability affect firms' actions. Heyes (1996) studies how making lenders liable for some part of the damages caused by their debtors affects both firms' cost of capital and level of care taken to avoid causing torts. Pitchford (1995) considers a similar question. Both conclude that the equilibrium cost of capital will (most likely) increase, but that the effect on firms' decision with respect to level of care is ambiguous. Yahya (1988) is closest in spirit to this paper, in that he allows firms to choose both a level of debt and a degree of care, and considers how the firm's decision changes under a variety of liability regimes.

By contrast, this paper seeks to make no recommendation as to how the legal system ought to be structured. Rather, the question posed here is to consider how the tort liability system, as constituted, will cause firms to respond to tort liability with changes in their capital structure.

2.3 Basic Model

2.3.1 Continuous firm returns framework

Consider a firm with the opportunity to pursue a one-period investment project. For simplicity, assume that the risk free interest rate is zero, and that the project's risk is entirely idiosyncratic. Investors are fully rational and risk neutral.[7]

Static tradeoff between cost of financial distress and tax savings

The firm must choose its time zero capital structure, which will be a combination of equity, with a time zero market value S_0, and one period debt with a promised time 1 payment D, which has a time zero market value B_0. Define V_0 as the sum of the time zero market value of the securities issued, $V_0 = S_0 + B_0$.

The project's terminal value is stochastic, and has a cumulative distribution function $G(V)$ and associated probability density function $g(V)$, with $G(V_{\min}) = 0$ and $G(V_{\max}) = 1$. The firm faces a tax rate τ on the time 1 payoff to equity , while debt holders' returns are not taxed. In the event that the realization of the project's terminal value is less than the face value of the debt, i.e. $V < D$, then the firm is bankrupt, and incurs fixed bankruptcy cost C. By assumption, $C \leq V_{\min}$.

[7] This is equivalent to assuming that the risk facing the firm is idiosyncratic, and investors are well diversified.

The market value of debt is given

$$B_0 = \int_{V_{\min}}^{D} (V - C)g(V)d(V) + \int_{D}^{V_{\max}} Dg(V)dV \tag{2.1}$$

while the market value of equity is given

$$S_0 = \int_{D}^{V_{\max}} (V - D)(1 - \tau)g(V)dV \tag{2.2}$$

The expected bankruptcy cost, C_0, is the cost of bankruptcy should it occur times the probability that the firm goes bankrupt. This can be expressed

$$C_0 = \int_{V_{\min}}^{D} Cg(V)dV \tag{2.3}$$

while the expected tax bill, T_0, is

$$\begin{aligned} T_0 &= \int_{D}^{V_{\max}} (V - D)\tau g(V)dV \\ &= \frac{\tau}{(1-\tau)} S_0 \end{aligned} \tag{2.4}$$

The expected value of the firm's cash flows is

$$\begin{aligned} E(V) &= \int_{V_{\min}}^{V_{\max}} Vg(V)dV \\ &= B_0 + S_0 + C_0 + T_0 \end{aligned} \tag{2.5}$$

in contrast to the market value of the firm's securities, V_0, which is

$$\begin{aligned} V_0 &= B_0 + S_0 \\ &= E(V) - C_0 - T_0 \end{aligned} \tag{2.6}$$

The firm's capital structure does not have an effect on the probability distribution governing the total cash flows to be shared between claimants. Therefore, the value maximizing level of D is that which minimizes the sum of expected bankruptcy costs and the expected tax bill. Since $\frac{\partial C_0}{\partial D} \geq 0$ and $\frac{\partial T_0}{\partial D} \leq 0$, there is a value D^*, $V_{\min} \leq D^* \leq V_{\max}$ that maximizes ex ante firm value V_0. This promised debt level occurs where $\frac{\partial C_0}{\partial D} = -\frac{\partial T_0}{\partial D}$.

Introduction of legal liability

Consider now the same firm, faced with the probability p that a tort litigant will appear, successfully sue the firm, and win a judgment J to be paid from the terminal asset value V. The claim has higher priority than equity, but lower priority than debt.[8] The expected payout to tort claimants (and the expected cost of tort liability), J_0, is given

$$J_0 = p\left(\int_{D+J}^{V_{\max}} Jg(V)dV + \int_{D+C}^{D+J} (V - D - C)g(V)dV\right) \tag{2.7}$$

The whole expression is multiplied by p, which is the probability that the plaintiffs win a judgment against the firm. The first term inside the brackets represents the range of terminal asset values where the firm is solvent, and must pay the tort claimants in full. The second term represents terminal asset values where the firm is bankrupt, and the tort claimants only collect their judgment after bankruptcy costs are paid and debtholders are paid fully. For asset values below $D + C$, tort claimants receive

[8] This is a simplification, in that in some jurisdictions other priority rules may apply.

nothing.

The market value of debt is

$$B_0 = \int_{V_{\min}}^{D} (V - C)g(V)dV + \int_{D}^{V_{\max}} Dg(V)dV + p \int_{D}^{D+C} (V - D - C)g(V)dV \tag{2.8}$$

The first two terms represent the market value of debt if there were no tort risk. The third term represents the impact the expected tort judgment has on the value of debt.[9]

The market value of equity is expressed

$$\begin{aligned} S_0 &= \int_{D}^{V_{\max}} (V - D)(1 - \tau)g(V)dV \\ &\quad -p\left(\int_{D}^{D+J} (V - D)(1 - \tau)g(V)dV + \int_{D+J}^{V_{\max}} J(1 - \tau)g(V)dV\right) \end{aligned} \tag{2.9}$$

The first term is a standard expression for the after-tax value of equity. The second term represents the expected cost to equity holders if the firm loses the tort judgment.[10]

Expected bankruptcy costs are

$$C_0 = \int_{V_{\min}}^{D} Cg(V)dV + p \int_{D}^{D+J} Cg(V)dV \tag{2.10}$$

[9]Since debtholders have priority, the size of the judgment, J, does not affect the value of debt. The effect is through the increased probability of bankruptcy in the event of losing the judgment. When tort claimants win the case, debtholders bear some portion of the bankruptcy costs for terminal asset values between D and $D + C$.

[10]Note that the expected transfer to tort claimants is also calculated on an after-tax basis.

The first term is the expected bankruptcy cost in the absence of tort liability, where the second term measures the expected increase in bankruptcy costs brought about by tort liability.

Finally, the expected tax bill is

$$\begin{aligned} T_0 &= (1-p)\int_{D}^{V_{\max}} (V-D)\tau g(V)dV + p\int_{D+J}^{V_{\max}} (V-D-J)\tau g(V)dV \\ &= \frac{\tau}{(1-\tau)} S_0 \end{aligned} \tag{2.11}$$

Once again, the market value of the securities the firm issues depends on the amount of debt issued:

$$\begin{aligned} V_0 &= S_0 + B_0 \\ &= E(V) - C_0 - T_0 - J_0 \end{aligned} \tag{2.12}$$

The goal of the firm is to set the debt level that maximizes the aggregate value of debt and equity. As can be seen from equation 2.12 , this is equivalent to setting the debt level, D^*, that minimizes the sum of expected bankruptcy, tax and tort judgment costs. The reformulated first order condition is therefore that the optimal debt level is chosen such that $\frac{\partial(C_0+T_0+J_0)}{\partial D} = 0$.

More detailed analysis depends on the distribution of firm asset returns. Comparative statics are unwieldly in the general case. As such, further analysis of how the optimal debt level changes is best conducted by studying specific distributional forms for the firm's asset returns.

Uniform distribution

Comparative statics are facilitated by making a distributional assumption regarding firm's asset returns. Assume that V is distributed uniformly between $V_{\min}$ and $V_{\max}$, i.e. that $g(V) = \frac{1}{V_{\max}-V_{\min}}$. The expected costs are then:

$$C_0 = \frac{C[D + pJ - V_{\min}]}{V_{\max} - V_{\min}} \tag{2.13}$$

$$J_0 = p\frac{\frac{1}{2}C^2 - J(C + D + \frac{1}{2}J - V_{\max})}{V_{\max} - V_{\min}} \tag{2.14}$$

$$T_0 = \tau\frac{\frac{1}{2}D^2 + pDJ + \frac{1}{2}pJ^2 - (D + pJ)V_{\max} + V_{\max}^2}{V_{\max} - V_{\min}} \tag{2.15}$$

Taking the derivative of the sum of the cost functions, setting to zero and solving for D yields

$$D^* = V_{\max} - \frac{C}{\tau} + \frac{1-\tau}{\tau}pJ \tag{2.16}$$

This can be compared to the firm's optimal debt level in the absence of tort liability, which is

$$D^*_{J=0} = V_{\max} - \frac{C}{\tau} \tag{2.17}$$

Since the tax rate, τ, is defined over $0 \leq \tau \leq 1$, the optimal debt level is increasing in the both the size of the judgment to be paid if the firm loses the case (J) and the probability of having to pay the judgment (p). So for the case where the firm's cash flows are distributed uniformly, an increase in the expected judgment pJ leads to an increase in the optimal face value of debt. This implies that the expected costs stemming from the increased probability of bankruptcy brought about by an increase in expected tort liability are outweighed by the asset shielding advantages of a relatively high debt level.

Importance of claim priority

The move away toward debt in the face of an increasing expected tort judgment highlights the importance of the priority of claims. In the above model, tort claimants collect only after debt holders have been paid. Tort claimants are, in effect, similar to involuntary subordinated debt holders.[11]

To see the importance the asset shielding effects of debt, consider the solution when *tort claimants* have priority. Here the firm still chooses a debt level to minimize the sum of J_0, C_0 and T_0; to reflect the change in priority, $\hat{J}_0 = pJ$. Firm returns are assumed to be uniformly distributed over $V_{\min}, V_{\max}$. For this specification, the optimal debt level is

$$\hat{D}^* = V_{\max} - \frac{C}{\tau} - pJ \tag{2.18}$$

With the asset shielding benefits of debt gone, the firm reduces its target debt level as the expected judgment increases.

This specification allows to consider a simple decomposition of the two effects on the optimal debt level that stem from an increase in judgment liability: the increase

[11] This analogy is only approximate. J represents the face value of the subordinated debtholders' claim, while, in a debt issue, it would be most likely that $p = 1$, as most debt issues require an attempt at repayment in all circumstances (lottery bonds and catastrophe bonds being exotic exceptions).

in expected bankruptcy costs, and the debt advantages of asset shielding. Since $\hat{D}^*$ is the optimal debt level in the presence of increased bankruptcy costs without the benefits of asset shielding, the bankruptcy cost effect can be defined as

$$\hat{D}^* - D^*_{J=0} = -pJ \tag{2.19}$$

The shift caused by the asset-shielding advantages of debt, when debt has higher priority, is then calculated

$$D^* - \hat{D}^* = \frac{1}{\tau}pJ \tag{2.20}$$

In the uniform distribution case, with $0 < \tau < 1$, the asset shielding advantages of debt outweigh the associated bankruptcy costs, and increases in tort liability lead to an increase in the optimal debt level.

Fraudulent conveyance

An important consideration for firms that choose to use capital structure as a defence against tort liability is whether or not their chosen capital structure will stand up to tort creditors' efforts to collect. As is true with any judgment proofing technique, there exists the risk that capital structure defences may be overturned. In this eventuality, the court would rule that setting a high debt level was done solely for the purpose of reducing the claim or tort creditors. The court would then be in a position to declare that the firm's capital structure amounts to a form of fraudulent conveyance, and award tort claimants higher than anticipated priority, thus rendering the firm's efforts to insulate its security holders from tort risk moot.

In the one-period model described above, the risk that the firm's defensive strategy would be overturned can be introduced relatively simply by assuming that whatever the firm's chosen capital structure, there exists the probability α that tort claimants will be awarded higher priority than debtholders. In this case, for a firm whose asset returns are uniformly distributed (as above), the firm's problem becomes to minimize the sum of

$$\begin{aligned} C_0 &= \frac{C[D + pJ - V_{\min}]}{V_{\max} - V_{\min}} && (2.21) \\ T_0 &= \tau \frac{\frac{1}{2}D^2 + pDJ + \frac{1}{2}pJ^2 - (D + pJ)V_{\max} + V_{\max}^2}{V_{\max} - V_{\min}} && (2.22) \\ J_0 &= \alpha pJ + p(1-\alpha)\frac{\frac{1}{2}C^2 - J(C + D + \frac{1}{2}J - V_{\max})}{V_{\max} - V_{\min}} && (2.23) \end{aligned}$$

The optimum debt level is

$$D^* = V_{\max} - \frac{C}{\tau} + \frac{1 - \tau - \alpha}{\tau} pJ \tag{2.24}$$

which implies that debt is decreasing in the probability α that the firm's defences will be overturned. This is not surprising, given that the expected asset shielding benefits of debt decrease in α, while the expected tax and bankruptcy costs remain the same.[12]

As the expected judgment cost pJ increases, changes in the optimum debt level

[12] The increase in the expected judgment cost comes at the expense of debtholders, the value of whose claims would fall should a court award tort claimants higher priority. Since debt is fairly priced, increases in α decrease the time 0 value of debt.

are no longer necessarily strictly increasing. It is readily apparent that

$$\frac{\partial D^*}{\partial(pJ)} = \frac{1-\tau-\alpha}{\tau} \tag{2.25}$$

implying that if the sum of the tax rate, τ, and the probability of claim priority being changed in favour of tort creditors, α, is greater than 1, then increases in the expected judgment result in the firm using less debt. The intuition is that for sufficiently high values of α, the expected asset shielding benefits of debt are reduced to the point that they are overtaken by the associated increased expected bankruptcy costs.

2.3.2 Discrete firm returns framework

An alternative specification is one where the firm's returns follow a discrete probability distribution. Consider the firm in the previous section. Instead of firm cash flows following a continuous distribution between $V_{\min}$ and $V_{\max}$, suppose that the cash flows follow a binomial distribution. At time 1, the firm's return is V_L with probability $(1-q)$ and V_H with probability q. All other variables are as defined in the previous section.

Tradeoff between financial distress and tax savings

I begin by reviewing the standard static tradeoff model. The market value of debt and equity, as well as expected bankruptcy costs and the expected tax bill, depend on the promised debt payment D. The market value of debt is

$$B_0 = \left\{ \begin{array}{cc} D & \text{for } D \leq V_L \\ (1-q)(V_L - C) + qD & \text{for } V_L < D \leq V_H \\ (1-q)V_L + qV_H - C & \text{for } V_H < D \end{array} \right\} \tag{2.26}$$

while the market value of equity is

$$S_0 = \left\{ \begin{array}{cc} ((1-q)V_L + qV_H - D)(1-\tau) & \text{for } D \leq V_L \\ q(V_H - D)(1-\tau) & \text{for } V_L < D \leq V_H \\ 0 & \text{for } V_H < D \end{array} \right\} \tag{2.27}$$

The expected tax bill is

$$T_0 = \frac{\tau}{1-\tau} S_0 \tag{2.28}$$

while expected bankruptcy costs are

$$C_0 = \left\{ \begin{array}{cc} 0 & \text{for } D \leq V_L \\ (1-q)C & \text{for } V_L < D \leq V_H \\ C & \text{for } V_H < D \end{array} \right\} \tag{2.29}$$

The expected value of the firm's cash flows is

$$\begin{aligned} E(V_1) &= (1-q_h)V_L + qV_H \\ &= B_0 + S_0 + T_0 + C_0 \end{aligned} \tag{2.30}$$

while the market value of the firm's securities is given

$$\begin{aligned} V_0 &= B_0 + S_0 \\ &= E(V_1) - T_0 - C_0 \end{aligned} \tag{2.31}$$

This implies that the market value of the firm is maximized when the term $(T_0 + C_0)$ is minimized. It is clear that for $D \leq V_H$, $\frac{\partial T_0}{\partial D} < 0$. However, since bankruptcy costs are fixed should they occur, and the probability of incurring these costs only increases at the debt levels $D = \{V_L, V_H\}$, the solution will be one of these two values. The marginal expected tax savings from moving from debt level $D = V_L$ to $D = V_H$ are

$$T_{0,V_L} - T_{0,V_H} = \tau q(V_H - V_l) \tag{2.32}$$

while the marginal expected bankruptcy costs are

$$(1-q)C \tag{2.33}$$

Define $\hat{C}$ as:

$$\hat{C} = \frac{q}{1-q}\tau(V_H - V_L) \tag{2.34}$$

Then for $C > \hat{C}$ the firm will set $D = V_L$, and for $C < \hat{C}$ the firm will set $D = V_H$.

Introduction of legal liability

Consider now the same firm, faced with the probability p that a tort litigant will appear, successfully sue the firm, and win a judgment J to be paid from the terminal asset value V, with priority higher than equity, but lower than debt holders. As in the previous section, the firm's problem is to set a debt level that maximizes the market value of its securities. However, unlike the previous section, the firm must additionally consider the risk posed by tort liability. On one hand, tort liability increases the risk of bankruptcy for a given level of debt, which would suggest a shift to a lower promised debt payment. On the other, where debt holders have priority over tort claimants, higher debt means a lower expected payment to tort claimants, suggesting a shift toward a higher promised debt payment. The net effect will depend on parameter values.

Once again, it is possible to consider a finite number of debt levels. For all debt levels $D < V_H$, a small increase of ε in D decreases the expected tax payment. However, if at debt level D, shifting to $D + \varepsilon$ does not lead to an increase in the expected bankruptcy cost, then D is not a potential solution. An ε increase in debt always leads to a decrease in the tax bill, and sometimes to a decrease in the expected judgment cost. Using this logic, one can easily show that the set of admissible debt levels which may solve the firm's problem are $D = \{V_L - J, V_L, V_H - J, V_H\}$.

Given the possible debt levels, and assuming $J < (V_H - V_L)$ and $V_L > C + J$, the expected cost of tort liability is

$$J_0 = \left\{ \begin{array}{cc} pJ & \text{for } D = V_L - J \\ pqJ & \text{for } D = V_L, V_H - J \\ 0 & \text{for } D = V_H \end{array} \right\} \tag{2.35}$$

Expected bankruptcy costs are

$$C_0 = \left\{ \begin{array}{cc} 0 & \text{for } D = V_L - J \\ p(1-q)C & \text{for } D = V_L \\ (1-q)C & \text{for } D = V_H - J \\ (1-q+pq)C & \text{for } D = V_H \end{array} \right\} \tag{2.36}$$

The market value of debt is

$$B_0 = \left\{ \begin{array}{cc} V_L - J & \text{for } D = V_L - J \\ V_L - p(1-q)C & \text{for } D = V_L \\ (1-q)(V_L - C) + q(V_H - J) & \text{for } D = V_H - J \\ (1-q)(V_L - C) + q(V_H - pC) & \text{for } D = V_H \end{array} \right\} \tag{2.37}$$

and the market value of equity is

$$S_0 = \left\{ \begin{array}{cc} (1-\tau)(q(V_H - V_L) + (1-p)J) & \text{for } D = V_L - J \\ (1-\tau)q(V_H - V_L - pJ) & \text{for } D = V_L \\ (1-\tau)q(1-p)J & \text{for } D = V_H - J \\ 0 & \text{for } D = V_H \end{array} \right\} \tag{2.38}$$

Tax is once again defined relative to equity,

$$T_0 = \frac{\tau}{1-\tau} S_0 \tag{2.39}$$

The solution to the firm's problem is to choose a promised debt payment from the set $D = \{V_L - J, V_L, V_H - J, V_H\}$ such that the sum of expected bankruptcy, tax and judgment costs are minimized. Define $C(D)$ as the total expected costs from choosing debt level D. The total expected costs from each of the four choices are

$$C(V_L - J) = q\tau(V_H - V_L) + [p + \tau(1-p)]J \tag{2.40}$$

$$C(V_L) = p(1-q)C + pq(1-\tau)J + q\tau(V_H - V_L) \tag{2.41}$$

$$C(V_H - J) = (1-q)C + q[p + \tau(1-p)]J \tag{2.42}$$

$$C(V_H) = [(1-q) + qp]C \tag{2.43}$$

As the firm moves progressively through to higher debt levels, the bankruptcy costs increase. To offset this effect, the expected cost of the tort judgment decreases (in bankruptcy states the firm doesn't pay), and expected tax costs decrease as well.

Depending on the parameter values, any of the four debt levels may prove to be optimal. Unlike the continuous case with a uniform distribution, the optimal debt level is not necessarily increasing in J. At low levels of J, firms are more likely to choose to accommodate the probability of facing a judgment by choosing either $V_L - J$ or $V_H - J$; as the potential cost of the judgment, J, increases, at some point the firm will no longer choose to accommodate the judgment, and will shift to either debt level V_L or V_H. If the potential judgment becomes sufficiently large relative to bankruptcy costs and other parameters, the firm will choose the maximum debt level, V_H.

The following diagram illustrates the potential for some firms to go through the entire range of possible debt choices depending on the level of the potential judgment, J. For the set of $\{V_L = 100, V_H = 200, p = 0.5, q = 0.5, \tau = 0.5, C = 65\}$, the optimal debt level D^* is on the vertical axis with the potential judgment, J, plotted on the horizontal axis:

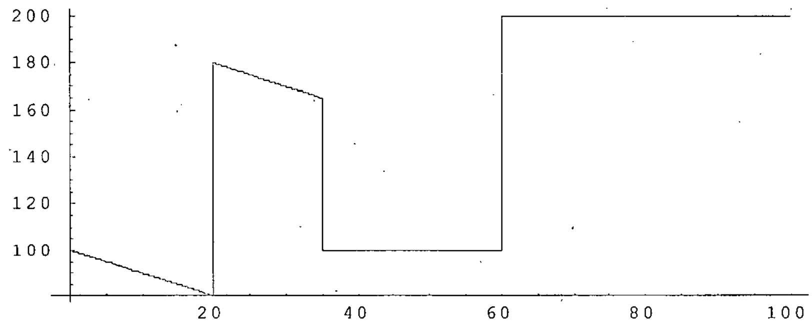

Figure 2.1: Choice of debt level for given sizes of judgment

In this example, in the absence of tort liability, the firm chooses V_L. As the potential judgment, J, increases, the firm initially chooses $V_L - J$, meaning that debt decreases dollar for dollar as J increases. At a critical point, the optimum jumps to $V_H - J$. As the potential judgment continues to increase, the firm eventually switches to a debt level of V_L. In this range, the optimum debt level is locally insensitive to changes in J. As J becomes sufficiently high, the firm moves to maximum debt, V_H.

The non-monotonicity arises in the discrete case because at low levels, small increases in the potential judgment do not warrant the increased risk of bankruptcy. However, as the size of the potential judgment rises relative to potential bankruptcy costs, eventually the asset-shielding benefits of debt outweigh the costs from bankruptcy, and the firm chooses to increase debt.

The discrete case highlights the importance of assumptions regarding asset returns. For different probability distributions of firm returns, the optimal capital structure response to changes in tort liability will differ.

2.4 The Impact of Liability Insurance

In some circumstances, firms may have the opportunity to buy liability insurance. Tillinghast Towers Perrin estimates that in 2003, over \$91 billion in tort costs were covered by firms' insurance policies. When liability insurance is an option the firm must make a joint capital structure-insurance coverage decision in order to maximize firm value. This section considers this decision problem.

2.4.1 Discrete firm returns with insurance

Assume that the structure of operating cash flows and tort liability is the same as in the previous discrete returns case. Now, the firm may choose to buy insurance, up to the value of the judgment, J, which pays off in the event that the firm loses a lawsuit and must pay a tort judgment. Assuming that the insurance is fairly priced, I dollars of coverage costs pI.

Proposition 1.1 *If a firm whose returns are binomially distributed chooses to insure, it will do so fully.*

Proof: The advantage of insurance is that it can be used to eliminate the probability that a judgment against the firm will cause it to incur bankruptcy costs. The disadvantage is that even under fairly priced insurance the premium is greater than the expected judgment cost, as long as the firm chooses a debt level such that there is some positive probability of bankruptcy. In the discrete case, buying anything less than full insurance does not decrease the probability of bankruptcy, while every dollar of coverage purchased *does* reduce the asset shielding advantage of debt. Therefore, if it is advantageous to buy the first dollar of insurance, it is more advantageous still to buy J dollars of insurance.[13] *QED*

In a setting with insurance, the firm has the same capital structure options as in the previous section, as well two new choices. The set is $D = \{V_L - J, V_{L,I=J}, V_L, V_H - J, V_{H,I=J}, V_H\}$, with $V_{L,I=J}$ and $V_{H,I=J}$ representing choices of debt level where the firm has chosen to fully insure against judgment liability.[14] The various costs of each choice are

$$J_0 = \left\{ \begin{array}{cc} pJ & \text{for } D = V_L - J \\ pqJ & \text{for } D = V_L, V_H - J \\ 0 & \text{for } D = V_{L,I=J}, V_{H,I=J}, V_H \end{array} \right\} \tag{2.44}$$

$$C_0 = \left\{ \begin{array}{cc} 0 & \text{for } D = V_L - J, V_{L,I=J} \\ p(1-q)C & \text{for } D = V_L \\ (1-q)C & \text{for } D = V_H - J, V_{H,I=J} \\ (1-q+pq)C & \text{for } D = V_H \end{array} \right\} \tag{2.45}$$

[13] See Huberman et al (1983).

[14] Assume that $V_{i,I=J}$ implies a promised payment to debtholders of $V_i - pJ$, and a promised payment to the insurer of pJ. Further assume that the insurer has priority, and $V_L - C > pJ$, guaranteeing that the insurer will be paid.

$$T_0 = \left\{ \begin{array}{c} \tau(q(V_H - V_L) + (1-p)J) \quad \text{for } D = V_L - J \\ \tau q(V_H - V_L) \quad \text{for } D = V_{L,I=J} \\ \tau q(V_H - V_L - pJ) \quad \text{for } D = V_L \\ \tau q(1-p)J \quad \text{for } D = V_H - J \\ 0 \quad \text{for } D = V_{H,I=J}, V_H \end{array} \right\} \tag{2.46}$$

as well as the insurance premium:

$$I_0 = \left\{ \begin{array}{cc} pJ & \text{for } D = V_{L,I=J}, V_{H,I=J} \\ 0 & \text{otherwise} \end{array} \right\} \tag{2.47}$$

Proposition 1.2 When insurance is available, $V_{L,I=J}$ dominates $V_L - J$.

Proof: Because a firm choosing $V_L - J$ never faces bankruptcy, it does not take advantage of the asset shielding effects of debt. For this firm, the cost of buying insurance is equal to the expected payment to judgment holders if uninsured. However, insurance allows the firm to take on a higher debt level, $D = V_L$, providing expected tax savings of τJ, without incurring bankruptcy risk. *QED*

Define the expected total costs for a given debt level D and insurance choice I as

$C(D, I)$. For the various combinations to be considerd, the expected total costs are

$$C(V_L, J) = pJ + \tau q(V_H - V_L) \tag{2.48}$$

$$C(V_L, 0) = p(1-q)C + pq(1-\tau)J + \tau q(V_H - V_L) \tag{2.49}$$

$$C(V_H - J, 0) = (1-q)C + q[p + \tau(1-p)]J \tag{2.50}$$

$$C(V_H, J) = (1-q)C + pJ \tag{2.51}$$

$$C(V_H, 0) = [(1-q) + qp]C \tag{2.52}$$

Depending on the parameter values, any of the five choices can be optimal. Of particular interest is the choice of $\{V_H, J\}$, where the firm chooses the high debt level but also purchases insurance. Firms making this decision are the only ones that "overpay" for insurance, to the extent that they surrender the asset shielding advantage of debt and pay for coverage in states where judgment holders would have been unable to collect. Despite this, it can still be an optimal decision if the tax savings brought about by being able to choose the high debt level V_H, without fear of increased bankruptcy risk brought about by tort liability, are sufficient.

However, firms will only ever consider one of $\{D, I\} = \{V_H - J, 0\}$ and $\{V_H, J\}$. Note that firms are indifferent between the two choices where

$$\frac{p}{(1-p)} = \tau \frac{q}{(1-q)} \tag{2.53}$$

The left side is the likelihood ratio of losing the tort judgment, and the right side is the likelihood ratio of realizing the high return multiplied by the tax rate. When the left side is greater than the right, implying relatively higher probability of losing the lawsuit, the firm will consider $\{V_H - J, 0\}$. When making this choice, the firm avoids overpaying the insurance premium, but accepts increased taxes when it realizes high returns and does not lose the lawsuit. When the right is greater the firm instead considers $\{V_H, J\}$. Here, the firm enjoys maximum tax savings, but at the cost of paying for insurance it does need when it realizes low terminal asset values. Note that which set the firm considers is independent of both J and C, meaning that firms will only ever consider one or the other.[15] Therefore there is no set of parameters $\{p, q, \tau\}$ where changes in J and C can produce as many as five different optimal debt choices.

In general, the availability of insurance reduces firms' propensity to accommodate a potential judgment by choosing either $D = V_L - J$ or $D = V_H - J$; the former is never chosen, and the latter considered only when the right side of equation 2.53 is greater than the left. This implies the existence of a greater number of states where the firm chooses a higher level of debt.

However, there also exist parameter values for which firms, who in the absence of insurance would have chosen V_L or V_H, shift to $\{V_L, J\}$ or $\{V_H, J\}$. This suggests that while the presence of insurance leads to more debt, it also leads to a greater number of states where tort creditors recover fully. Insurance leads to more firms with deep pockets.[16] To the extent that tort judgments are legitimate attempts to redress those who have been harmed in some way, this is a socially desirable effect.

[15] Of course, how the cost of either $\{V_H - J, 0\}$ or $\{V_H, J\}$ compare to the costs of the other three options depends critically on the relative values of C and J.

[16] Strictly speaking, the deep pockets belong to the insurance companies with whom the firm has contracted. From the plaintiff's perspective, this distinction is not important.

2.5 Joint and Several Liability

An essential extension of the analysis is to consider how a firm's behaviour changes when it is dependent on the outcome for other firms. This situation arises where firms are jointly and severally liable for a given tort. In the simplest example of how this type of liability works, a plaintiff sues two defendants who both contributed to causing her harm. If a judgment is found in the plaintiff's favour, each defendant is ordered to pay half the judgment. However, if one of the defendants becomes insolvent, the other becomes responsible for whatever remaining portion of the judgment needs to be paid.

This extension is far from being an esoteric detail. Joint and several liability is now standard for many types of torts in many juriscitions. According to the 2004 report by Tillinghast Towers Perrin, "there appears to be a shift in the types of liabilities that make up the total tort costs in the U.S., from individuals suing individual entities to groups of plaintiffs taking legal action against one or more entities".

In this circumstance, capital structure choice is the outcome of a game. The capital structure choices of a set of firms who share liability for a given tort become interdependent. The externalities among firms give rise to the possibility of multiple equilibria. The implication is that firms' decisions with respect to capital structure now depend on the decisions of other firms: specifically, those of their co-defendants. As is demonstrated below, circumstances can emerge where a firm would choose a relatively conservative debt level, if it knew that its potential co-defendants would do the same, thereby committing to being solvent to able pay their share of the potential judgment. However, if the co-defendants choose higher debt levels, implying that the firm would be left on its own to cover the cost of the entire judgment, the firm's decision would change; it too would shift to a higher debt level.

2.5.1 Known returns case

First, consider a case with two firms, each of whose asset value will be V at the end of the period. At that point, the firms will lose a tort case with probability p, in which case they will be jointly and severally liable for paying the judgment J. If both firms are solvent, they each owe $\frac{J}{2}$ to the tort claimants. Should one firm not be able to pay the judgment, then the other is responsible for the full amount.

The firms each choose a debt level, D_i. It is straigtforward to show that each firm's optimal debt level will always be one of the three values $\{V - J, V - \frac{J}{2}, V\}$.[17] Without loss of generality, I restrict the analysis to these three values.

The optimal level for each firm depends not only on the parameter values for $\{p, \tau, C, J\}$, but also on the other firm's choice of debt level. This occurs because each firm must consider whether or not the other firm has chosen a capital structure

[17] The tax and asset shielding advantages to debt financing are continually increasing in D, while the expected bankruptcy costs only jump at specific debt levels. The critical levels at which an epsilon increase in debt will (sometimes) increase the probability of bankruptcy are $V - J$ and $V - \frac{J}{2}$. V is the highest possible debt level, as there are no asset shielding or tax benefits to choosing a debt level beyond this point.

that will leave it solvent and able to pay its share of the judgment should the firms lose the lawsuit. In terms of impact on the other firm, the choices $V - J$ and $V - \frac{J}{2}$ can be grouped together, since both of these levels leave the firm able to meet its share of the obligation. However, choosing a debt level of V imposes an externality on the other firm; should the case be lost, the other firm will be faced with a bill for the full judgment.

Expected costs from each choice of debt level must be calculated based on the other firm's decision. A cost function $C_i(D_i, D_j)$ is defined as the combined expected bankruptcy, tax and tort judgment costs for firm i, given that firm i chooses debt level D_i and firm j choose debt level D_j. The cost functions to be considered are:

$$\begin{aligned}
C_i(V - J, \{V - J, V - \frac{J}{2}\}) &= \frac{J}{2}\left[p + \tau + (1 - p)\tau\right] \\
C_i(V - J, V) &= J\left[p + (1 - p)\tau\right] \\
C_i(V - \frac{J}{2}, \{V - J, V - \frac{J}{2}\}) &= \frac{J}{2}\left[p + (1 - p)\tau\right] \\
C_i(V - \frac{J}{2}, V) &= \frac{J}{2}\left[p + (1 - p)\tau\right] + pC \\
C_i(V, \{V - J, V - \frac{J}{2}, V\}) &= pC
\end{aligned}$$

Looking at the cost functions, it is immediately apparent that $D_i = V - \frac{J}{2}$ is never a best response when the other firm sets $D_j = V$; it is dominated by a symmetrical response of $D_i = V$. It is also clear that $D_i = V - J$ in response to $D_j = \{V - J, V - \frac{J}{2}\}$ is dominated by $D_i = V - \frac{J}{2}$. Therefore, $\{D_i, D_j\} = \{V - J, V - J\}, \{V - J, V - \frac{J}{2}\}, \{V - \frac{J}{2}, V\}$ are not possible equilibria.

The remaining possible equilibria are $\{D_i, D_j\} = \{V - \frac{J}{2}, V - \frac{J}{2}\}, \{V, V\}$, and $\{V - J, V\}$. Which equilibrium will prevail depend on the size of the possible judgment J in relation to the other variables. When

$$J < C\frac{p}{p + (1-p)\tau}$$

$\{D_i, D_j\} = \{V - \frac{J}{2}, V - \frac{J}{2}\}$ is the only possible equilibrium. When

$$J > 2C\frac{p}{p + (1-p)\tau}$$

$\{D_i, D_j\} = \{V, V\}$ is the only possibility. However, in the region

$$C\frac{p}{p + (1-p)\tau} \leq J \leq 2C\frac{p}{p + (1-p)\tau}$$

either of the two equilibria is possible. While it is not difficult to show that the firms would prefer the $\{D_i, D_j\} = \{V - \frac{J}{2}, V - \frac{J}{2}\}$ equilibrium, the firms do not necessarily have the opportunity to choose. Once one firm has adopted the high capital structure, the other must follow suit, and neither will have an incentive to deviate.

A simple diagram illustrates this point, for parameter values $\{V = 10, p = .5, \tau = .5, C = 3\}$:

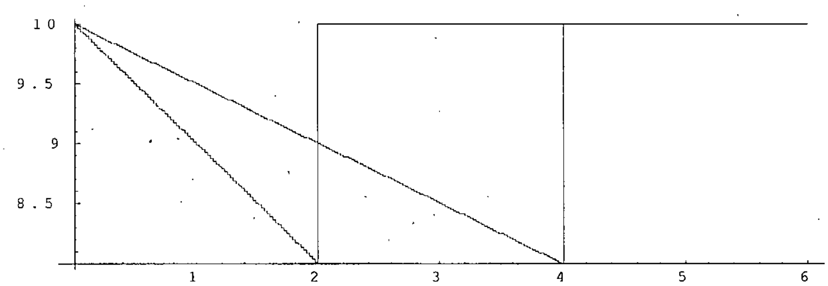

Figure 2.2: Joint and several liability and known returns

Values of J are plotted on the bottom axis, while debt levels are on the vertical axis. The two curves represent optimal responses, depending on the other firm's choice. The curve which is initially more steeply sloped represents optimal choices of D_i when the other firm's debt level is $D_j = V$. The curve that is initially less steeply sloped represents optimal choices of D_i when the other firm's debt level is $D_j = \{V - J, V - \frac{J}{2}\}$. In the region $J < 2$, the optimal debt level for both firms is $V - \frac{J}{2}$. When $J > 4$, both firms choose V. However, for $2 \leq J \leq 4$, two equilibria are possible. Either $\{D_i, D_j\} = \{V - \frac{J}{2}, V - \frac{J}{2}\}$, or $\{D_i, D_j\} = \{V, V\}$.

2.5.2 Binomially distributed returns case

Consider two firms, identical to those described previously. Instead of facing a certain return V, each firm faces symmetric,independently binomially distributed returns. That is, each either returns V_L or V_H at time 1, and each has the same probability of realizing a high return, $q_i = q_j = q$. The firms will be ordered to pay a judgment J with probability p. Again, the firms are jointly and severally liable.

The firms each choose a debt level, $\{D_i, D_j\}$. Again, each firm need only consider a finite number of potential debt levels. The initial set to be considered is $D = \{V_L - J, V_L - \frac{J}{2}, V_L, V_H - J, V_H - \frac{J}{2}, V_H\}$. The complication comes when each firm must consider the expected bankruptcy, tax and tort judgment costs associated with each debt level in response to the possible choices of the other firm. An equilibrium is a situation where each firm's choice is a best response to the other firm's debt level, recognizing that each firm's asset returns are stochastic.

The expected costs associated with firm i's decision to choose debt level D_i, given that firm j chooses D_j, are defined as $C_i(D_i, D_j)$. The cost functions are as follows:

$$\begin{aligned} & C_i(V_L - J, \{V_L - J, V_L - \frac{J}{2}\}) \\ = \;& q\tau\,(V_H - V_L) + \frac{J}{2}\,(p + \tau + (1-p)\tau) \end{aligned} \tag{2.54}$$

$$\begin{aligned} & C_i(V_L - J, \{V_L, V_H - J, V_H - \frac{J}{2}\}) \\ = \;& q\tau\,(V_H - V_L) + \frac{J}{2}\,(p(2+q)(1-\tau) + \tau) \end{aligned} \tag{2.55}$$

$$\begin{aligned} & C_i(V_L - J, V_H) \\ = \;& q\tau\,(V_H - V_L) + J\,(p + (1-p)\tau) \end{aligned} \tag{2.56}$$

$$\begin{aligned} & C_i(V_L - \frac{J}{2}, \{V_L, V_H - J, V_H - \frac{J}{2}\}) \\ = \;& q\tau(V_H - V_L) + p(1-q)^2 C + \frac{J}{2} p\tau(\frac{1}{\tau} + \frac{1}{p} - 1 - (1-q)q) \end{aligned} \tag{2.57}$$

$$\begin{aligned} & C_i(V_L - \frac{J}{2}, \{V_L - J, V_L - \frac{J}{2}\}) \\ = \;& q\tau\,(V_H - V_L) + J\left(\frac{p}{2}(1-\tau) + \tau\right) \end{aligned} \tag{2.58}$$

$$\begin{aligned} & C_i(V_L - \frac{J}{2}, V_H) \\ = \;& q\tau(V_H - V_L) + p(1-q)C + \frac{J}{2}\,(p(1-\tau) + \tau(1-pq)) \end{aligned} \tag{2.59}$$

$$
\begin{aligned}
& C_i(V_L, \{V_L - J, V_L - \frac{J}{2}\}) \\
= \quad & q\tau(V_H - V_L) + p(1-q)C + \frac{J}{2}pq(1-\tau)
\end{aligned}
\tag{2.60}
$$

$$
\begin{aligned}
& C_i(V_L, \{V_L, V_H - J, V_H - \frac{J}{2}\}) \\
= \quad & q\tau(V_H - V_L) + p(1-q)C + \frac{J}{2}pq(1-\tau)(2-q)
\end{aligned}
\tag{2.61}
$$

$$
\begin{aligned}
& C_i(V_L, V_H) \\
= \quad & q\tau(V_H - V_L) + p(1-q)C + (1-\tau)pqJ
\end{aligned}
\tag{2.62}
$$

$$
\begin{aligned}
& C_i(V_H - J, \{V_L - J, V_L - \frac{J}{2}\}) \\
= \quad & (1-q)C + \frac{J}{2}\left(pq(1-\tau) + 2\tau q\right)
\end{aligned}
\tag{2.63}
$$

$$
\begin{aligned}
& C_i(V_H - J, \{V_L, V_H - J, V_H - \frac{J}{2}\}) \\
= \quad & (1-q)C + \frac{J}{2}\left(pq(1-\tau)(2-q) + 2\tau q\right)
\end{aligned}
\tag{2.64}
$$

$$
\begin{aligned}
& C_i(V_H - J, V_H) \\
= \quad & (1-q)C + Jq(p + \tau(1-p))
\end{aligned}
\tag{2.65}
$$

$$
\begin{aligned}
& C_i(V_H - \frac{J}{2}, \{V_L - J, V_L - \frac{J}{2}\}) \\
= \quad & (1-q)C + \frac{J}{2}q(p + \tau(1-p))
\end{aligned}
\tag{2.66}
$$

$$
\begin{aligned}
& C_i(V_H - \frac{J}{2}, \{V_L, V_H - J, V_H - \frac{J}{2}\}) \\
= \quad & (1-q)(1+pq)C + Jq(p + \tau(1-p))
\end{aligned}
\tag{2.67}
$$

$$C_i(V_H - \frac{J}{2}, V_H)$$
$$= (1-(1-p)q)C + Jq(p+\tau(1-p)) \tag{2.68}$$

$$C_i(V_H, D_j)$$
$$= (1-(1-p)q)C \tag{2.69}$$

While this set of cost functions is difficult to analyze analytically, it is possible to do some numerical experimentation. For several sets of parameter values, it becomes clear that while all of the debt levels may be optimal in some circumstances, it is generally true that equilibria involve firms choosing to match each other's debt level; the relevant cost functions to consider are then defined by equations 2.54, 2.58, 2.61, 2.64, 2.67 and 2.69.[18]

Further, it is apparent that both firms choosing the debt level $V_L - J$ will not be an equilibrium. Provided that both firms choose $V_L - J$, each has an incentive to move to $V_L - \frac{J}{2}$. Since neither set of choices ever results in either firm going bankrupt, the firms prefer $V_L - \frac{J}{2}$, as the higher debt level provides a lower expected tax bill, with expected bankruptcy costs and judgment costs remaining unchanged.

The optimal decision depends on the parameter values p, q, C, V_L, V_H, τ and J. Comparisons are probably most relevant when made as follows.

[18]It may be possible that two firms will choose different debt levels, with the higher debt firm's choice imposing a greater externality on the other firm. However, for this to be an equilibrium, it would have to be the case that $C_i(H,H) - C_i(H,L) \geq C_i(L,H) - C_i(L,L)$. This condition can be interpreted as being that a shift to the higher debt level has a greater negative impact on the other firm when the other firm is already at the higher debt level. There is not any evidence that a set of parameters meeting this condition, and being otherwise consistent with the model, exists.

For the sake of exposition, I examine the capital structure decision of two identical firms, to be held jointly and severally liable for the amount J should they lose the court case, with the parameters $\{V_L = 50, V_H = 100, p = .5, q_i = .5, \tau = .5, C_i = 10\}$ held constant, as the aggregate amount of the potential judgment J varies. I compare this with the debt decision taken by one firm, faced with the same potential liability J, with parameters $\{V_L = 100, V_H = 200, p = .5, q = .5, \tau = .5, C = 20\}$. This is relevant because aggregate "industry" revenues and bankruptcy costs are the same as for the two smaller firms, as is the potential judgment.

As J varies (values on the horizontal axis), the two firms' choice of debt level is plotted:

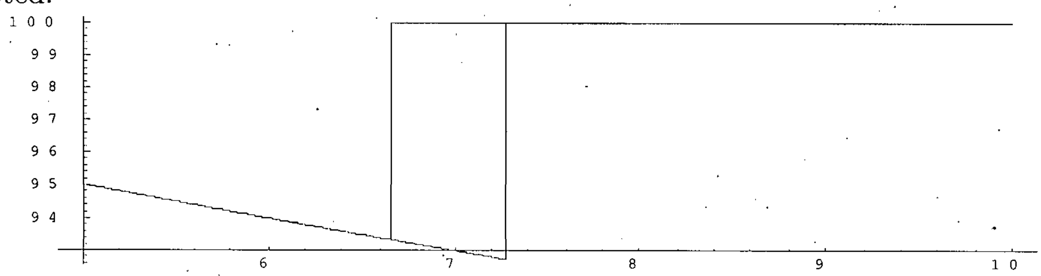

Figure 2.3: Joint and several liability and binomially distributed returns

In the range $0 \leq J < 6.67$, the equilibrium is $D_i = D_j = V_H - J$. For $J > 7.27$ the equilibrium is the maximum possible debt level, $D_i = D_j = V_H$. For the range $6.67 \leq J \leq 7.27$, either of the other two equilbria are possible; the firms either both select $V_H - J$, or they both select V_H.

The decision of the larger single firm, faced with the entire liability itself, is plotted:

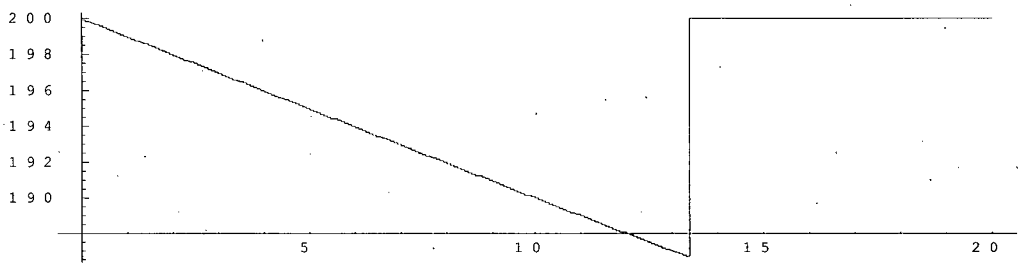

Figure 2.4: Single large firm base case debt level

Here, the firm initially chooses debt level $V_H - J$, and shifts to V_H when $J = 13.33$.

In the example, it's apparent that the effect of joint and several liability on the capital structure decision is ambiguous. At low levels of J, the two firms choose to accommodate not only their own initial share of the judgment, but also that of the other firm, recognizing that their co-defendant could go bankrupt. For values of $J < 6.67$, tort claimants' expected recovery is higher than the one-firm case.[19]

For $J > 7.27$, both firms will shift to the highest possible debt level, V_H, and tort claimants' expected recovery drops to zero. In the single firm case, this shift does not occur until $J \geq 13.33$. For values of J such that $6.67 \leq J \leq 7.27$, two equilbria are possible. The firms will either both accommodate the full share of the judgment by setting $D_i = D_j = V_H - J$, or both firms will shift to the highest possible debt level V_H.

Generally, joint and several liability serves to reduce debt levels and increase

[19]In the one firm case, the single firm has realize the high asset return for the tort claimants to be able to collect the judgment should they win. The probability of their having a claim against a solvent defendant, conditional on having won the case, is q. On the other hand, when faced with two defendants each choosing a debt level of $V_H - J$, only one need be solvent. The probability of being able to collect is $q(2-q)$, which is greater than q.

expected tort claimant recovery at low judgment levels, while it decreases expected recovery at higher debt levels. When the cost of losing a judgment is low relative to the costs associated with going bankrupt, both firms choose a capital structure which would allow them to pay should they realize high returns. Tort claimants end up benefitting from a "diversification" effect. Rather than being exposed to the risk that a single defendant's deep pockets will be emptied by the vagaries of business risks, defendants have two entities to pursue, and enjoy the increased probability that at least one's pockets will remain deep. As the potential size of the judgment increases, however, the defendants start to impose externalities on each other. To protect themselves from having to pay the other's share of the judgment, both choose aggressive capital structures to insulate themselves against the potential judgment. Aggregate debt therefore tends to increase, and tort claimants recover in fewer states of the world than they would against a single, larger defendant.

2.6 Two Period Model

An extension which adds some richness to the analysis is to consider how firms will make capital structure decisions as tort liability evolves over time. The anticipated risk of losing a major lawsuit is not static. As new information emerges about the likelihood that the firm has caused a tort against another party, or about the magnitude of the harm caused, all market participants will reasonably update their expectations about the probability of the firm having to pay tort claimants.

By the same token, neither is capital structure static. Firms have the flexibility to increase or decrease their debt level as time goes on, continually trading off the asset and tax shielding benefits of debt versus the expected cost of financial distress.

By considering a two-period model, it is possible to consider how a firm's capital structure changes through in time, in response to changes in tort risk.

2.6.1 Basic two period model

Consider a firm pursuing a project with a two period life. At the end of the first period, the firm receives an update as to the distribution governing the project's final distribution. At the end of the second period, the terminal asset value is realized. The distribution of the time 1 reported asset value is governed by probability density function $g(V_1)$ and cumulative distribution function $G(V_1)$, with $G(V_{1\,\min}) = 0$ and $G(V_{1\,\max}) = 1$. The realization of V_1, which is the signal received at time 1, is the time 1 expectation of the eventual time 2 cash flow realization. This time 2 cash flow realization, V_2, is distributed with probability density function $g(V_2)$ and cumulative distribution function $G(V_2)$, with $G(V_{2\,\min}) = 0$ and $G(V_{2\,\max}) = 1$. $V_{2\,\min}$ and $V_{2\,\max}$ are defined such that $E_1(V_2 \mid V_1) = V_1$. This implies that $E_0(V_1) = E_0(V_2)$.[20]

The firm finances its operations through a mix of one-period debt and equity. At

[20] An example of the type of situation this set of distributions is meant to describe would be as follows. A firm receives an updated signal, V_1, uniformly distributed between 50 and 150. The firm's eventual value, V_2, will be uniformly distributed between $V_1 - 50$ and $V_1 + 50$. As such, $E_1(V_2) = V_1$, while the time zero expectations are $E_0(V_1) = E_0(V_2) = 100$.

time zero, the firm issues debt with a face value of D_1, payable at time 1. At time 1, after receiving the updated signal V_1, the firm chooses a face value of debt for the second period, D_2, payable at time 2.

Taxes are payable at both time 1 and time 2, as a fraction of asset value at that date. Debt payments shield assets from the tax collector, so the taxes payable at a given date t are $T_t = \tau(V_t - D_t)$ whenever $V_t > D_t$, zero otherwise.

Finally, bankruptcy costs C are incurred at either date whenever $V_t < D_t$. This can be interpreted as costly renegotiation.

By assumption costs, there are no costs associated with adjusting capital structure. In the absence of any frictions, the firm chooses a debt level at date $t-1$ such that minimizes $E_{t-1}[C_t + T_t]$. Because the relative distributions at both dates between $V_{t\,\mathrm{min}}$ and $V_{t\,\mathrm{max}}$ are the same, the firm chooses the same *relative* debt level each period, denoted D_t^*.[21]

2.6.2 Two period model with tort liability

Consider now a firm faced with tort risk, such that there is some chance that they will pay a judgment J at time 2. The time zero risk of having to pay the judgment is $\frac{p}{2}$. At time 1, the firm receives an updated signal about the lawsuit's prospects.

[21] Of course, the absolute second period debt level will be higher when $V_1 > E_0(V_1)$, and lower when $V_1 < E_0(V_1)$.

With probability $\frac{1}{2}$ the suit is found to have no merit (lawsuit risk falls to zero), and with probability $\frac{1}{2}$ the plaintiffs' chances of winning the suit improve to p.

By assumption, the firm's time 1 tax charge does not change, and is still defined as $T_1 = \tau(V_1 - D_1)$. However, the risk of financial distress increases, as security holders take the time 2 judgment risk into account when valuing their claims. The time zero expectation of the time 1 bankruptcy cost becomes

$$E_0(C_1) = C\left[\frac{1}{2}\int_{V_{1\,\min}}^{D_1} g(V_1)dV_1 + \frac{1}{2}\int_{V_{1\,\min}}^{D_1+pE_1(J)} g(V_1)dV_1\right]$$

The firm solves for the promised time zero debt payment which minimizes $E_0[C_1 + T_1]$. Because tort risk increases the probability bankruptcy, and period 1 debt shields assets from the tax collector only, the optimum debt level $\hat{D}_1^*$ is lower than the optimum level in the case where expected tort liability is zero at time 1.

At time 1, if the risk of tort liability disappears, the optimum time 2 promised debt payment is D_2^*, the same as the optimal level in the case without tort liability. On the other hand, if the probability of having to pay a judgment increases to p, the firm chooses $\hat{D}_2^*$ so as to minimize the sum of $E_1[C_2 + T_2 + J_2]$. This is essentially the same problem as defined in the single period model described earlier in this paper.

Whether or not $\hat{D}_2^*$ is a higher relative debt level than $\hat{D}_1^*$ depends on distributional assumptions about the distribution of V_i. As shown in the one period model, however, if V_i is uniformly distributed, then the debt level is increasing in expected tort liability.

Since this is the case where expected tort liability increases from time zero to time 1, for a uniform distribution it will be the case that $\hat{D}_2^* > \hat{D}_1^*$.

This leads to an interesting conclusion. For at least some distributional assumptions about firm returns, the further resolution of uncertainty about tort risk leads to an increase in the firm's relative debt level, whether that resolution increases or decreases tort risk.

While this is somewhat counterintuitive, it can be explained as follows. At time zero, the probability of having to pay a judgment is $p/2$. For a given time 1 face value of debt, this risk increases the expected bankruptcy cost, providing an incentive to lower debt. However, because the judgment, if eventually paid, will only be paid at time 2, the promised time 1 debt payment does not provide any asset shielding benefits.

At time 1, there are two possible resolutions of uncertainty about tort risk. In the case where tort liability disappears, the probability of bankruptcy for a given relative debt decreases. Since the tax shielding benefits of debt don't change, the new optimum relative debt level is higher than it was at time 0. In the case where tort liability increases, the expected bankruptcy costs also increase. However, promising a debt payment at time 2 helps shield the firm's security holders from tort creditors, which is not true of the promised time 1 debt payment. For distributions where the asset shielding benefits of debt outweigh the increased expected bankruptcy costs, even increased tort liability at time 1 will lead to the firm choosing to increase its debt level.

2.7 Conclusion

This paper seeks to explain how tort liability will affect a firm's optimal capital structure. While other papers have made the point that limited liability will affect economic agents' incentives with respect to tort risk, very few have sought to endogenize the firm's decision about in which states it will be solvent.

A key characteristic of tort risk is that its impact on cash flows available to security holders depends on the structure of security holders' claims. Put another way, capital structure matters greatly when determining the potential expense payable to tort claimants.

Recognizing this, firms with exposure to tort liability will have an incentive to adjust capital structure to respond optimally. The lower creditor priority of tort claimants implies two effects when debt and tort risk interact. The first is that tort liability brings about an increased probability of bankruptcy. Where this effect predominates, the firm will choose to move away from debt. The second effect is that debt provides an asset shielding advantage, preserving cash flow rights for the firm's debt holders at the expense of tort claimants. Where this effect is dominant, increased tort risk will cause the firm to choose more debt.

I specify two simple models to examine the interaction of these effects, one where firm returns are distributed continuously over an interval, and another where firm returns are distributed binomially. The different results from these two illustrations demonstrate the importance of assumptions regarding firm cash flows. Depending on the nature of the firm's returns, and the values of the various input parameters, either the bankruptcy effect or asset shielding effect can dominate.

I also consider how liability insurance affects the outcome. Fairly priced liability insurance is in effect overpriced for any firm with positive probability of bankruptcy, due to the asset shielding effects of debt. However, the model in this paper demonstrates that there are circumstances where firms will still choose to purchase insurance. The model also indicates that the availability of insurance can lead to greater amounts of debt being issued, at the same time as providing tort litigants with deeper-pocketed targets.

I test how firms' capital structure decisions change when several smaller firms are jointly and severally liable for a judgment, and compare their behaviour to that of a larger entity faced with the same potential judgment. I find that for relatively low tort amounts, debt levels tend to be lower, and tort claimants' expected recovery greater. However, for higher judgment amounts, debt levels tend to increase, and tort claimants will expect to recover less.

Finally, I examine how the capital structure decision changes as liability evolves through time. In the model, I find that the resolution of uncertainty about tort risk leads to an increase in the debt level, whether the resolution is one of lower tort risk or higher tort risk.

Tort liability is a major source of risk for firms today. I have explained why it is unique, and why firms must consider its unique properties when determining the optimal capital structure. Empirical work studying how firms do adjust their capital structure to address changes in tort risk is a potentially fruitful avenue for future research.

2.8 Bibliography

[1] Avila-Nores, Matias and Stephen L. Schwarcz, "Ring-Fencing and Other Bankruptcy-Remote Techniques", Duke University School of Law Working Paper, 2002.

[2] Bergmann, Karyn S., "Bankruptcy, Limited Liability and CERCLA: Closing the Loophole and Parting the Veil", University of Maryland School of Law Working Paper Series No. 2004-02, 2004.

[3] Carroll, Stephen J., Deborah Hensler, Allan Abrahamse, Jennifer Gross, Michelle White, Scott Ashwood and Elizabeth Sloss, "Asbestos Litigation Costs and Compensation", RAND Institute for Civil Justice Documented Briefing, 2002.

[4] Dahiya, Sandeep and David Yermack, "Wealth Creation and Destruction from Brooke Group's Tobacco Litigation Strategy", New York University Law School Working Paper No. CLB-00-007, 1999.

[5] Delaney, Kevin J., "Power, Intercorporate Networks, and 'Strategic Bankruptcy'", *Law & Society Review*, 1989, Vol. 23, No. 4, 643-666.

[6] Fleming, John G., "Mass Torts", *The American Journal of Comparative Law*, Summer 1994, Vol. 42, No. 3, 507-529.

[7] Hansmann, Henry and Reinier Kraakman, "Toward Unlimited Shareholder Liability for Corporate Torts", *The Yale Law Journal,* May 1991, Vol. 100, No. 7, 1879-1934.

[8] Heyes, Anthony G., "Lender Penalty for Environmental Damage and the Equilibrium Cost of Capital", *Economica*, May 1996, Vol. 63, No. 250, 311-323.

[9] Huberman, Gur, David Mayers and Clifford W. Smith, Jr., "Optimal Insurance Policy Indemnity Schedules", *The Bell Journal of Economics*, Autumn 1983, Vol. 14, No. 2, 415-426.

[10] Kornhauser, Lewis A. and Richard L. Revesz, "Apportioning Damages Among Potentially Insolvent Actors", *The Journal of Legal Studies*, June 1990, Vol. 19 No. 2, The Law and Economics of Risk, 617-651.

[11] Kraus, Alan and Robert H. Litzenberger, "A State-Preference Model of Optimal Financial Leverage", *The Journal of Finance*, September 1973, Vol. 28, No. 4, 911-922.

[12] Lopucki, Lynn M., "Virtual Judgment Proofing: A Rejoinder", *The Yale Law Journal,* March 1998, Vol. 107, No. 5, 1413-1434.

[13] Malani, Anup and Charles Mullin, "The Effect of Joint and Several Liability

on the Bankruptcy Rate of Defendants: Evidence from Asbestos Litigation", University of Virginia Law School Working Paper, 2004.

[14] Painter, Christopher M.E., "Tort Creditor Priority in the Secured Credit System: Asbestos Times, the Worst of Times", *Stanford Law Review*, April 1984, Vol. 36, No. 4, 1045-1085.

[15] Pitchford, Rohan, "How Liable Should a Lender Be? The Case of Judgment-Proof Firms and Environmental Risk", *The American Economic Review*, December 1995, Vol. 85, No. 5, 1171-1186.

[16] Priest, George L., "Understanding the Liability Crisis", *Proceedings of the Academy of Political Science*, Vol. 37, No. 1, 196-211.

[17] Roe, Mark J., "Corporate Strategic Reaction to Mass Tort", *Virginia Law Review*, February 1986, Vol. 72, No. 1, 1-59.

[18] Ringleb, Al H. and Steven N. Wiggins, "Liability and Large-Scale, Long-Term Hazards", *The Journal of Political Economy,* June 1990, Vol. 98, No. 3, 574-595.

[19] Schwartz, Alan, "Products Liability, Corporate Structure, and Bankruptcy: Toxic Substances and the Remote Risk Relationship", *The Journal of Legal Studies*, December 1985, Vol. 14, No. 3, Critical Issues in Tort Law Reform: A Search for Principles, 689-736.

[20] Shavell, Steven, "The Judgment Proof Problem", *International Review of Law and Economics,* 1986, Vol. 6, 45-58.

[21] Shavell, Steven, "Minimum asset requirements and compulsory liability insurance as solutions to the judgment-proof problem", *RAND Journal of Economics,* Spring 2005, Vol. 36, No. 1, 63-77.

[22] Swiss Re, "Sigma: The economics of liability losses - insuring a moving target", 2004, No. 6.

[23] Tillinghast - Towers Perrin, "U.S. Tort Costs: 2004 Update".

[24] Tillinghast - Towers Perrin, "U.S. Tort Costs and Cross-Border Perspectives: 2005 Update", 2006.

[25] Tung, Frederick, "Taking Future Claims Seriously: Future Claims and Successor Liability in Bankruptcy", Working Paper, 1998.

[26] Vairo, Georgene, "Mass Torts Bankruptcies: The Who, The Why and The How", Loyola Law School Research Paper No. 2003-21, 2003.

[27] White, James J., "Corporate Judgment Proofing: A Response to Lynn LoPucki's The Death of Liability", *The Yale Law Journal,* March 1998, Vol. 107, No. 5, 1363-1412.

[28] Yahya, Moin A., "Bankruptcy and Torts", Working Paper, 1998.

Chapter 3

The Capital Structure of Insurers: Theory and Evidence

3.1 Introduction

In the simplest economic models of insurance markets, which ignore transactions costs of any kind, risks are priced at actuarially fair values. This prediction depends on one of two sets of assumptions: the *pooling* theory of insurance assumes that insured risks are independently distributed and large in number; the *transfer* theory of insurance assumes that risks are independent of aggregate wealth in the economy and can be transferred through the issuance of equity to a perfect capital market (Marshall (1976)). Recent research in insurance economics has shown that the observed dynamics of insurance premiums and contracts can be explained only by a failure of both sets of assumptions. Aggregate uncertainty, combined with imperfections in the equity market, can disrupt the transfer of risks to the capital market in ways that explain insurance market dynamics (e.g., Grøn (1994), Winter (1988,1994)). This connection is not surprising, since imperfections of some sort are necessary to explain even the existence of insurance intermediaries. The empirical tests in this recent literature have focussed on time series implications of insurance pricing and capital flows.

This paper explores the cross-sectional variation in insurers' capital structures: the choice by stock insurers of the mix of equity and liability.[1] As in the standard theory of optimal capital structure in finance, predictions of the theory must rely on specific capital market imperfections. We focus here on the simplest one: that issuing and maintaining additional equity is costly. Our model yields testable implications with a focus (appropriate for an analysis of insurance markets) on the liability side of the market.

Section 2 of this paper develops the simplest model of an insurance market with costly equity, in a two-period setting. For equity to have any role in an insurance

[1]The capital structure decision for insurers, being a financial intermediary, is different from the decision faced by non-financial firms. Non-financial firms have some underlying assets which generate cash flows; the capital structure decision relates to how to finance those assets by apportioning claims to cash flows between debt and equity holders.

Insurance companies' liabilities are the insurance policies themselves, whice arise naturally in the course of doing business. The question of how much equity to maintain relates to what kind of "cushion" the firm requires to credibly back the policies it issues. This problem is similar to the one made by banks; faced with a given level of deposits (liabilities), banks must determine how much equity it requires to maintain capital adequacy.

market there must be aggregate uncertainty, or dependence among insured risks; the absence of a law of large numbers means that equity is necessary to back up promises to pay claims in the event of adverse realizations of aggregate shocks.[2] Accordingly, the key comparative static issue that we focus on is the impact of increasing aggregate uncertainty. We consider separately the cases of aggregate uncertainty in the *loss* incurred conditional upon an accident and uncertainty in the *probability* of an accident (i.e. dependence among the *events* of individual accidents). In the former case, the total equity issued by a competitive insurance market is increasing in the degree of uncertainty (and linear in a parameterized example). In the latter case, equity may be increasing then decreasing as a function of uncertainty. In both cases, the *ratio* of equity to revenue is increasing in uncertainty.

Section 3 tests the theory using cross-sectional data on U.S. property-liability insurers. While the theory is developed for competitive markets, by assuming that each insurer is operating in a different set of one or more competitive markets, we can use firm-level data in the tests. The focus is on tests of two hypotheses. The first is the implication of the static model, that leverage is decreasing in aggregate uncertainty. The second is an implication of previous dynamic models of competitive insurance markets (Grøn (1994) and Winter (1994)) that external equity is more costly than internal equity – specifically that there is a positive cost to the "round-

[2]Aggregate uncertainty is necessary, that is, in the limit as the number of consumers gets large. With independence, the law of large numbers would allow the risk of bankruptcy to be avoided by a vanishingly small amount of equity per policy.

trip" of distributing an amount of cash then raising the same amount in external equity. Previous tests of this implication focus on the time series behavior of insurance premiums. The empirical analysis here is complementary, based not on prices but directly on capital structure decisions. The paper also offers a link between the recent insurance market literature and corresponding empirical results in tests of capital structure for non-financial corporations: Titman and Wessels (1988) and Rajan and Zingales (1994) find negative relationships between leverage and past profitability; an explicit dynamic theory and tests are offered by Fischer, Heinkel and Zechner (1989).

3.2 The Optimal Capital Structure of Insurers

We describe the capital structure choice of an insurance firm in the simplest possible model. The key assumption must be that risks are dependent, i.e. subject to aggregate uncertainty or common factors. We consider separately the cases of dependence in the *events* of accidents and dependence in the *size* of losses incurred.

3.2.1 Aggregate Uncertainty in Accident Losses

Assumptions

We consider a competitive market for insurance. On the demand side of the market, a large number of individuals each face with a known probability p the loss of wealth. The size of the loss is itself random, taking on the value H with probability

λ and L with probability $(1-\lambda)$. If the risks faced by individuals were *independently* distributed then – given a large number of individuals – insurance would be provided at a fair premium with no need for equity. The optimal capital structure would (in the limit) have zero equity. We introduce a role for equity by assuming that the random losses faced are dependent among individuals. In fact, for simplicity, losses are identical for those experiencing an accident. In short, each individual faces a two-stage lottery, with "accident - no accident" in the first stage and "L or H" in the second. Across individuals the first stage outcomes are independently distributed, while the second-stage outcomes are identical.

The individuals are expected utility maximizers and the gain from exchange in the insurance market arises because they are risk averse. We take the simple case of identical individuals, with initial wealth W and utility U, where $U' > 0$ and $U'' < 0$.

Ex ante, a large number of stock insurers issue equity and then issue insurance policies. An insurance policy is assumed to be *non-participating.* That is, the contract with any individual specifies a payment that is contingent only on the individual's loss experience: I_L dollars if the individual experiences a loss of L, I_H dollars with a loss of H. The premium is denoted by P. We constrain the insurance contracts and equity to satisfy a limited liability constraint, so that the contracts promised by the insurer must be credibly backed by the equity issued. We denote by E the equity per policy issued. A second constraint is that the promised payment in any accident state cannot exceed the accident loss in that state. This can be justified by a moral hazard assumption that an individual has the ability to cause an accident intentionally.

In a perfect capital market, the cost of issuing and maintaining equity would be zero. Equityholders would be indifferent between investing through the insurance corporation and investing through their personal portfolios. It is evident that in reality equity cannot be issued by an insurer and maintained without limit at zero cost. The costs include agency costs of having corporate management intermediate between investment in assets and shareholders; the administrative costs of issuing equity; the signalling costs of issuing equity and the double-taxation of corporate income.[3] We do not model these costs explicitly, but simply assume that equity cannot be raised at zero cost. Specifically, we assume that it costs $(1+c)E$ dollars to raise E dollars of equity, which is returned to claimants on the firm's assets. The term cE represents the net cost of maintaining equity.

Equityholders price equity according to the expected value of net payments that they are to receive; this reflects an assumption that the uncertainty in losses, while not diversifiable in the insurance market, is diversifiable in the stock market. Interest rates are zero. The supply of insurance is taken to be competitive, which means that any capital structure E and policy (P, I_L, I_H) consistent with zero expected return to equityholders will be supplied if it is demanded. On the demand side, the individuals

[3]It is evident that means of distributing cash to shareholders other than by dividends cannot be relied upon as costless alternatives.

choose the most preferred policy among policies offered by the market.

This model yields, as an equilibrium, the choice of an insurance policy that maximizes the expected utility of the individual among all the policies yielding zero expected return to stockholders. The issue of concern is how the equilibrium values of equity and the structure of liabilities vary with uncertainty in losses.

Remarks

This is the simplest model within which we can address the impact of dependence in risks and costly equity capital structure decisions. Several features of the simple model abstract from reality. First, we have taken the *form* of the insurance contract, the nonparticipating contract, as exogenous. This can be justified formally with an assumption that an individual can verify only his own accident experience. It includes the simplification that no mutual insurance is available. Second, in this static model we do not capture any distinction between the costs of maintaining equity, and the costs of adjusting equity. The evidence from the recent literature is that this distinction is important for explaining the dynamics of pricing and capital flows. In the empirical section, we shall in fact offer some evidence of the cost advantage of internal capital – and, implicitly, of the value of extending this model to a dynamic context.

Equilibrium

Consider first the payoffs to equityholders and individual demanders of insurance, under the contract $[E, P, I_L, I_H]$ when this contract is offered to all individuals.[4] The payoffs to an individual who does not experience an accident is $W - P$. The payoff to an individual who experiences an accident with loss X, for $X = L$ or H is $W - P - X + I_X$. The net payoff to equityholders (per policy issued) in the event that the common accident loss is X is $-cE + P - pX$, since a proportion p of individuals experience an accident.

The contract offered in a competitive insurance market will maximize expected utility subject to three constraints. The first is a limited liability constraint, that the payment to accident victims in each event X must not exceed the sum of internal equity, $(1 - c)E + P$. That is, $pI_X \leq (1 - c)E + P$. The second is a participation constraint for insurers, that the expected profit be non-negative: $-cE + P - p[\lambda I_H + (1 - \lambda)I_L] \geq 0$. The third is the constraint that $I_X \leq X$. The following results are easily proved.

Proposition 3.1: *If $c = 0$, then the equilibrium insurance policy involves full coverage of each loss.*

Lemma 3.1: *With $c > 0$, the participation constraint is binding and:*

(a) *the constraint $I_L \leq L$ is binding: $I_L = L$.*

[4] It is convenient to consider the equity E as one component of the contract; it backs the promise to pay the claims I_L and I_H.

(b) *the limited liability constraint is binding in the event* $H : pI_H = (1-c)E + P$

(c) $I_H < H$

Proposition 1 is the standard perfect capital market benchmark. Lemma 1 is for the case of $c > 0$. Here, without the "moral hazard" constraint that $I_L \leq L$, low losses would actually be *more* than fully covered.[5] The lemma allows us to simplify the contract specification and payoffs: A contract can without loss of generality now be described as a pair (P, E). Individuals receive a net payoff of $W - P$ in any event except a high-loss accident, and in the event of a high-loss accident they receive $W - P - H + (P + (1-c)E)/p = W - H + (\frac{1-p}{p})P + \frac{1-c}{p}E$. The *gross* payout to shareholders is zero in the event of high accident losses (where the limited liability constraint is binding), so that the expected profit to shareholders from issuing a contract (P, E) is

$$-cE + P - p\left[\lambda I_H + (1-\lambda)I_L\right]$$

Using Lemma 1 (a) and (b), this expected profit can be written

$$-(c + \lambda(1-c))E + (1-\lambda)P - p(1-\lambda)L$$

In sum, we can characterize the equilibrium insurance contract as the solution to the following problem:

$$\max_{E,P} \quad (1-p\lambda)U(W-P) + p\lambda U\left(W - H + \frac{1-p}{p}P + \frac{1-c}{p}E\right) \tag{3.1}$$

[5]This result follows because the events of an accident are independent across individuals, and therefore the market offers wealth transfers between the events of "accident" and "no accident" at an actuarially fair rate. The individual optimum therefore requires the equality of marginal utility in the event of no-accident and expected marginal utility conditional upon an accident. To achieve this equality, since high losses are not fully covered, low losses must be more than fully covered.

subject to

$$-(c+\lambda(1-c))E+(1-\lambda)P-p(1-\lambda)L=0 \tag{3.2}$$

Letting the multiplier on the constraint be μ, the first order conditions with respect to E and P respectively are:

$$\lambda(1-c)U'(\cdot)-(c+\lambda(1-c))\mu=0 \tag{3.3}$$

$$-(1-p\lambda)U'(W-P)+\lambda(1-p)U'(\cdot)+(1-\lambda)\mu=0 \tag{3.4}$$

where $U'(\cdot)=U'\left(W-H+\frac{1-p}{p}P+\frac{1-c}{p}E\right)$. Solving 3.3 for μ and substituting into 3.4, we obtain

$$-(1-p\lambda)U'(W-P)+\lambda\left(1-p+\frac{(1-\lambda)(1-c)}{c+\lambda(1-c)}\right)U'(\cdot)=0 \tag{3.5}$$

Equations (3.2) and (3.5) characterize the optimal contract.[6]

Our interest is in the impact on the equilibrium contract of an increase in aggregate uncertainty. We represent an "increase in uncertainty" as a mean preserving spread in the conditional distribution of losses, but with the further restriction that λ remains constant in this increase. That is, an increase in aggregate uncertainty is represented as $dH>0$ with the restriction $dL=-\lambda/(1-\lambda)\cdot dH$. Totally differentiating (3.2) and (3.5) with this substitution yields

$$-(c+\lambda(1-c))dE+(1-\lambda)dP+\lambda p dH=0 \tag{3.6}$$

[6] Note that if c equals zero, so that we have a perfect capital market, then (3.5) implies that the two marginal utilities are equal, which in turn implies full insurance. This equation shows also that if (3.5) is positive, then the coverage is less than full in the bad state.

$$\frac{\lambda(1-c)}{p}\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)U''(\cdot)dE \tag{3.7}$$
$$+\left((1-p\lambda)\,U''\,(W-P)+\lambda\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)\frac{1-p}{p}U''(\cdot)\right)dP$$
$$-\lambda\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)U''(\cdot)dH=0$$

From (3.6) and (3.7) it can be shown that

$$\begin{matrix} dE/dH \\ \\ dP/dH \end{matrix} = A^{-1}\begin{bmatrix} \lambda\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)U''(\cdot) \\ \\ -\lambda p \end{bmatrix} \tag{3.8}$$

where A is given by

$$\begin{bmatrix} \frac{\lambda(1-c)}{p}\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)U''(\cdot) & (1-p\lambda)U''(W-P)+\lambda\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)(\frac{1-p}{p})U''(\cdot) \\ -(\lambda(1-c)+c) & 1-\lambda \end{bmatrix}$$

Proposition 3.2: *With aggregate uncertainty in the size of losses, an increase in uncertainty leads to*

(a) *an increase in equity, E;*

(b) *an increase in the premium, P; and*

(c) *an increase in the equity-to-premium ratio, E/P.*

Proof: We can write (3.8) in shorthand as $\begin{bmatrix} dE/dH \\ dP/dH \end{bmatrix} = \begin{bmatrix} -a & -b \\ -g & d \end{bmatrix}^{-1} \begin{bmatrix} -e \\ -f \end{bmatrix}$ with all of the lower-case letters on the right-hand side positive. (This can be shown using $U'' < 0$.) Solving for dE/dH gives $dE/dH = \frac{1}{-(ad+bg)}(-de-bf) > 0$, proving (a). Solving for dP/dH yields $dP/dH = \frac{1}{-(ad+bg)}(-ge+af)$. Substituting back in the terms for $(-ge+af)$ yields

$$\begin{aligned}(-ge+af) &= (c+\lambda(1-c))\left[\lambda\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)U''(\cdot)\right] \\ &\quad -\frac{\lambda(1-c)}{p}\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)\lambda pU''(\cdot) \\ &= \left(c+\lambda(1-c)-\lambda^2(1-c)\right)\left(1-p+\frac{1-\lambda(1-c)}{c+\lambda(1-c)}\right)U''(\cdot)\end{aligned}$$

Therefore $(-ge+af)<0$, hence $dP/dH>0$, proving (b). To prove (c), re-write (3.2) as

$$\frac{P}{E}=\frac{c+\lambda(1-c)}{1-\lambda}+p\frac{L}{E}$$

from which we have

$$\frac{d\left(\frac{P}{E}\right)}{dH}=\frac{p}{E^2}\left[E\frac{dL}{dH}-L\frac{dE}{dH}\right]=\frac{p}{E^2}\left[-E\frac{\lambda}{1-\lambda}-L\frac{dE}{dH}\right]$$

which is negative since $dE/dH>0$ by (b). Leverage is therefore decreasing in uncertainty. *QED*

3.2.2 Uncertainty in Accident Probabilities

Assumptions

The alternative structure is one in which common factors are in the events of accidents. We assume now that the loss from an accident is known, and equal to L, but, because of dependence in the events of accidents, the frequency of accidents is random. This frequency, p, is assumed to take on two possible values, a and b, with

$b > a$. The term λ now represents the probability of the frequency b of accidents. The *ex ante* probability of an accident for any individual is $\bar{p} = (1 - \lambda)a + \lambda b$.

A contract now involves the promise of a payment I in the event of an individual accident in exchange for the premium P. In contrast to the case of uncertain losses that are identical across individuals, where contractual promises for cash flows are always met, we must introduce here the notion of bankruptcy. An insurer with equity-per-contract E is bankrupt if $P + E - pI < 0$. We allow for the possibility that bankruptcy involves the loss of specific assets, interpreted as a reputation for prudence, or other bankruptcy costs. As before, issuing equity requires a transaction cost of c per unit.

As before, we consider the contract offered by a competitive market to identical, risk-averse consumers. This is the contract that maximizes individual expected utility subject to a zero-profit constraint.

Equilibrium

Depending on the market parameters, especially the size of bankruptcy costs and λ, the equilibrium may or may not involve bankruptcy in the event that the accident frequency is b. In the case where bankruptcy costs are sufficiently large, the equilibrium contract in this model will satisfy the solvency constraint in both states. We consider this a reasonable approximation, in light of the regulatory solvency constraints faced by firms. These constraints do not, evidently, reduce the probability of bankruptcy to zero; but the rate of bankruptcy is very small with less than one percent of policies defaulted on in any year. In understanding the costs and benefits in the choice of an equity ratio by a firm facing existing solvency regulation, and generating testable implications regarding this choice, approximating the regulation as a complete constraint against bankruptcy is useful.[7]

The expected net profit to shareholders from the policy (P, I) with equity E is $-cE + P - [(1-\lambda)a + \lambda b]\, I$. When the firm is subject to a no-bankruptcy condition for both events, a and b, the gross return to shareholders in the event b is zero, since (it is easily shown) excess equity will not be issued. The amount of equity, E, will be chosen given the contract (P, I) to meet the no-bankruptcy constraint in event b: that the amount of equity remaining after payment of the costs cE, covers net losses: $(1-c)E + P \geq bI$.

Let $\bar{p} \equiv (1-\lambda)a + \lambda b$. The equilibrium contract is characterized by the maximum of expected utility subject to the no-bankruptcy constraint and the zero profit constraint:

$$\max_{I,P,E}(1-\bar{p})U(W-P) + \bar{p}U(W-P-L+I) \tag{3.9}$$

subject to

$$(1-c)E + P \geq bI \tag{3.10}$$

$$-cE + P - [(1-\lambda)a + \lambda b]\, I = 0 \tag{3.11}$$

[7] We have elsewhere considered in more detail the effect of actual solvency regulation on insurance markets (Winter (1991)).

Proposition 3.3: *In the case of uncertain probabilities with a no-bankruptcy constraint,*

(a) *an increase in uncertainty leads to an increase in E if uncertainty is sufficiently small. That is, $dE/db|_{\bar{p}} > 0$ if $b - a$ is sufficiently small.*

(b) *For larger levels of uncertainty, dE/db may be positive or negative.*

(c) *For utility satisfying $U'(W) > 0, U''(W) < 0$, optimal equity is increasing in the degree of risk aversion $\gamma = -\frac{U''(W)}{U'(W)}$.*

Proof:

Solving the first (no-bankruptcy) constraint for I yields

$$I = \frac{(1-c)E + P}{b} \tag{3.12}$$

and substituting it into the zero profit constraint yields

$$-\left[c + \frac{\bar{p}}{b}(1-c)\right] E + \left(1 - \frac{\bar{p}}{b}\right) P = 0 \tag{3.13}$$

Let

$$A \equiv \frac{c + \frac{\bar{p}}{b}(1-c)}{1 - \frac{\bar{p}}{b}} \tag{3.14}$$

Solving (3.13) for P and substituting this value into the objective function in (3.9) yields the following as a characterization of the optimal equity:

$$\max_E (1-\bar{p})U(W - AE) + \bar{p}U\left(W - L + \left[\frac{1-c}{b} + (\frac{1}{b} - 1)A\right] E\right) \tag{3.15}$$

Letting W_{NA} and W_A be shorthand for the realized wealth in the no-accident and accident states respectively, and setting the derivative of this expression with respect to E to zero gives

$$-(1-\bar{p})AU'(W_{NA}) + \bar{p}\left[\frac{1-c}{b} + (\frac{1}{b} - 1)A\right]U'(W_A) = 0 \tag{3.16}$$

This is the first order condition. Since $\lim_{b \to \bar{p}} A = 0$, this derivative is unbounded as $b \to \bar{p}$. Therefore, for b sufficiently close to $\bar{p}$ (i.e. for sufficiently small uncertainty) the optimal E is positive. However, for zero uncertainty ($b = \bar{p}$), the optimal contract is easily shown to yield the standard full insurance solution: $I = L$, $P = \bar{p}L$ and $E = 0$. It follows that, holding $\bar{p}$ constant, $dE/db > 0$ for b sufficiently close to $\bar{p}$.

To prove (b), a parametric example suffices. Define the form of the utility function as $U(W) = -e^{-\gamma W}$, and define parameter values of $\{\bar{p}, c, \gamma, L, W\} = \{.1, .1, 1, .25, 1\}$. Optimal equity of E^*, defined as being the value of E which satisfies (3.16), can be evaluated as b changes. The following diagram plots optimal equity given changes in b as it ranges from .101 to .300:

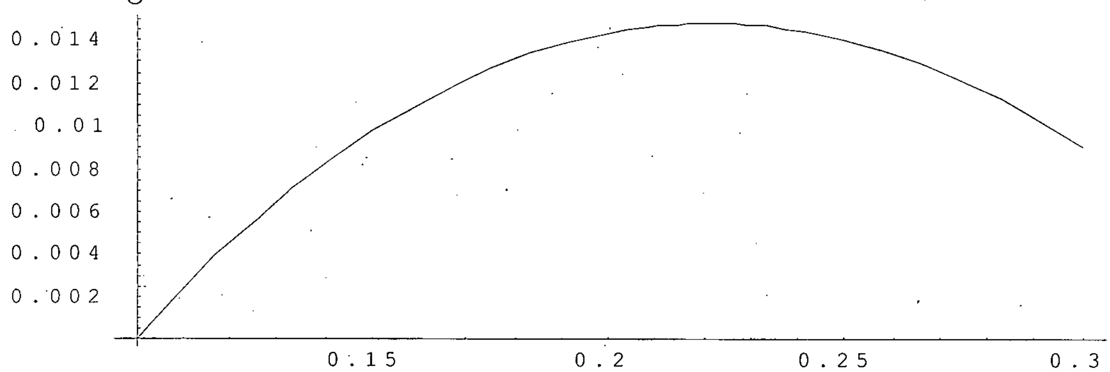

Figure 3.1 - Optimal equity under changing uncertainty

Initially, E^* increases as b increases, reaching a maximum where $b = 0.2214$. Beyond this point, E^* is decreasing as uncertainty in the probability of loss increases further.

To prove (c), consider the first order condition, (3.16). First, note that because $U'(W) > 0$ and $U''(W) < 0$, it is true that $U'(x) < U'(y)$ where $x > y$. It can be shown that $(1-p)A > \bar{p}\left[\frac{1-c}{b} + (\frac{1}{b} - 1)A\right]$. For (3.16) to hold, it must be the case that $U'(W_{NA}) < U'(W_A)$, meaning that $W_{NA} > W_A$. Wealth is greater in the no-accident state, which means that in equilibrium there is less than full insurance.

Given that there is less than full insurance, consider how increasing risk aversion affects the optimal level of equity, E^*. Define the measure of risk aversion as $\gamma = -\frac{U''(W)}{U'(W)}$. As noted in the previous paragraph, it is true that, $x > y$, $U'(y) - U'(x) > 0$. It is further the case that $\frac{\partial[U'(y)-U'(x)]}{\partial\gamma} > 0$. One can define $(\tilde{W}_{NA}, \tilde{W}_A)$ as satisfying (3.16) for some level of risk aversion $\tilde{\gamma}$, and $(\hat{W}_{NA}, \hat{W}_A)$ as satisfying (3.16) for a higher level of risk aversion $\hat{\gamma}$. It must be the case that

$$\hat{W}_{NA} - \hat{W}_A < \tilde{W}_{NA} - \tilde{W}_A \qquad (3.17)$$

since the higher degree of risk aversion $\hat{\gamma}$, combined with the requirement than (3.16) hold, implies that the difference between the no-accident and accident wealth levels must be less than for the lower risk aversion case. Put another way, as risk aversion increases, the optimal contract moves closer to full insurance. Substituting into (3.17) for $\{\hat{W}_{NA}, \hat{W}_A, \tilde{W}_{NA}, \tilde{W}_A\}$, and defining $\{\hat{E}^*, \tilde{E}^*\}$ as the optimal levels of equity for the two different levels of risk aversion, we have

$$\left(W - A\hat{E}^*\right) - \left(W - L + \left[\frac{1-c}{b} + \left(\frac{1}{b} - 1\right) A\right] \hat{E}^*\right) \tag{3.18}$$

$$< \left(W - A\tilde{E}^*\right) - \left(W - L + \left[\frac{1-c}{b} + \left(\frac{1}{b} - 1\right) A\right] \tilde{E}^*\right) \tag{3.19}$$

(3.19) can be simplified:

$$\left(A + \left[\frac{1-c}{b} + \left(\frac{1}{b} - 1\right) A\right]\right) \left(\tilde{E}^* - \hat{E}^*\right) < 0 \tag{3.20}$$

The left term of (3.20) is positive, meaning that $\left(\hat{E}^* - \tilde{E}^*\right)$ is negative. Therefore, optimal equity for the higher degree of risk aversion, $\hat{E}^*$, is greater, meaning that optimal equity is increasing in the degree of risk aversion. *QED*

At the heart of the comparative statics in proposition 3 are two off-setting effects of an increase in uncertainty on equity. Holding constant the amount of coverage issued, I, an increase uncertainty b implies an increase in the value of equity, E, that is necessary to cover the claims at a given premium. This is the *input effect.* The amount of coverage will drop, however, as a consequence of the higher cost of offering any amount of coverage; this feeds back to a decrease in E: an *output effect.* When uncertainty is sufficiently low, the input effect dominates and when uncertainty is high, the output effect may dominate. The two effects can be seen in the total differentiation of the no-bankruptcy condition, which yields $dE/db = \frac{I+b(dI/db)-dP/db}{1-c}$. The first two terms of this are, respectively, the input effect and the output effect. Endogenizing the change in P through total differentiation of (3.11), holding $\bar{p}$ constant, yields

$$\frac{dE}{db} = I + (b - \bar{p})\frac{dI}{db}$$

again showing a decomposition into the input and the output effects.

Proposition 4: In the case of uncertain probabilities with high bankruptcy costs, an increase in uncertainty leads to an decrease in the leverage ratio P/E.

Proof: Solve (3.13) for $\frac{P}{E} = A$, from which

$$\frac{\partial\left(\frac{P}{E}\right)}{\partial b} = -\frac{\bar{p}}{(b - \bar{p})^2}$$

The ratio of liabilities to equity is therefore decreasing in uncertainty. *QED*

To summarize the main comparative static results that flow from the model: increasing aggregate uncertainty leads to an increase in optimal amount of equity when the uncertainty is in the size of the loss (conditional upon an accident) but a non-monotonic relationship in the case where the uncertainty (i.e. dependence) is in the events of accidents. In both cases, the equity to premium ratio is increasing in uncertainty. Since the premium is the market valuation of an insurer's liability (this liability being of course the promise of insurance payouts), the inverse of this ratio is the liability to equity ratio, analogous to the debt-equity ratio conventionally used to summarize capital structure. The negative relationship between the liablity-equity ratio and the level of uncertainty is the first implication of the theory that we will test in the next section of this paper.

Extension: Initial Equity Endowment

An additional implication follows from a simple extension of the model. Let E^* represent the optimal equity that would be issued in a competitive insurance market under either set of assumptions that we have set out in the model. Suppose now that firms are *endowed with* an amount of internal equity, A, at the beginning of the period on which they do not have to incur issuance costs. This endowment represents internal capital inherited from retained profits earned previously. (We retain the static model assumption that the equity of the firm is distributed entirely to shareholders at the end of the period.) Then the equilibrium amount of equity, $\widehat{E}$, in the extended model is $\widehat{E} = \max(E^*, A)$: if $A \leq E^*$, the equilibrium will be identical to the equilibrium analyzed in the model above. Those firms endowed with substantial internal equity will earn rents on this endowment, but the contract will reflect the opportunity cost of capital at the margin, including the issuance costs. For these values of A, $d\widehat{E}/dA = 0$. On the other hand, when A exceeds E^*, then the entire equity is retained until the end of the period and $d\widehat{E}/dA = 1$.[8] It is clear that over the region $A \geq E^*$, the premium P is non-increasing in A; P is decreasing in A over the subset of this region where the limited liability constraint is binding. It therefore follows that over this region $d(\widehat{E}/P)/dA > 0$, or $d(P/\widehat{E})/dA > 0$. The prediction is that past profitability should have a negative impact on leverage.

[8] Allowing firms to distribute internal equity (through a special dividend or share repurchase) at the beginning of the period does not change the essential results.

3.3 Evidence

3.3.1 Introduction

Cross-sectional data on a sample of U.S. property-liability insurers allows us in this section to provide evidence on two aspects of insurers' capital structure decisions. The first is the implication from our model that insurers' leverage should be decreasing in the uncertainty faced in predicting average risks. While the equilibrium equity in our model is for particular cases non-monotonic in uncertainty, leverage – as measured by the ratio of insurance revenue to equity – is unambiguously decreasing in uncertainty.

The second aspect of capital structure on which we offer evidence is the relative costs of internal and external equity. The extension of our model to include an endowment of low-cost equity suggests that firms with greater access to less costly internal capital will use less leverage. Recent theory on the economic dynamics of insurance markets relies on the assumption that internal capital is less costly than external equity. By a cost advantage to internal capital, we mean simply that there is a positive cost to the round-trip of distributing an amount of cash to equityholders, then raising the same amount through the issuance of new equity. (The basis for such a cost is well-developed in the literature, e.g. Myers and Majluf (1984)). Up to now, this assumption has been tested for insurance markets using the time series of insurance market pricing. The implication of this assumption for the cross-section is that leverage should be decreasing in recent profitability, since this profitability leads to greater accumulation of internal equity.[9]

3.3.2 Empirical Proxies and Estimation

The firm specific data are collected from A.M. Best's Aggregates and Averages annual reports on consolidated property-casualty insurance companies. These statutory financial information are filed by insurance companies to National Association of Insurance Commissioners (NAIC) to assist insurance commissioners in regulating and monitoring insurance companies licensed in their respective state. The selected sample covers 852 U.S. property-casualty stock insurance companies from 1999 to 2004.

	Mean	Std. Dev.	Minimum	Maximum
NPE/E	1.0455	0.6814	0.0003	4.5491
SD	0.1946	0.4310	0.0097	4.8975
SIZE	1,147,510	3,837,997	769	50,959,623
PROFIT	0.0491	0.2080	-4.0868	1.5836

Table 3.1: Descriptive Statistics

The cross-sectional regressions of firm capital structure (leverage) on three hypothesized determinants— uncertainty of insurance loss, firm size, and past profitability, are specified as

[9]The tests of both hypotheses for insurance markets are parallel to tests of capital structure hypotheses for general corporations that have been offered in the financial economics literature (e.g., Bradley, Jarrell and Kim (1984), Titman and Wessels (1988)).

$$\log(NPE/E) = \alpha + \beta \cdot \log(SD) + \delta \cdot \log(SIZE) + \eta \cdot PROFIT + \varepsilon \tag{3.21}$$

where *NPE* is Net Premiums Earned, *E* is the Policyholders Surplus, *SD* is the uncertainty of the insurance loss, *SIZE* is the firm size, and *PROFIT* is past profitability. These empirical proxies are defined as follows.

• Policyholders' Surplus E: the equity of a property-casualty insurance firm.

• Net Premium Earned *NPE*: the total insurance policy revenue from policies issued during a given year, adjusted for any increase or decrease in liabilities for unearned premiums during the year.[10]

• Loss Ratio: the ratio of incurred losses and loss adjustment expenses to net premium earned.

• Capital Structure (NPE/E): is measured as the ratio of Net Premium Earned (*NPE*) to Policyholders' Surplus (E) in 2004 This reflects the relationship between the current volume of net insurance liability and the equity.

• Uncertainty of the insurance loss SD: is represented by the standard deviation of the loss ratio from 1999 to 2004. The theoretical model predicts an inverse relationship between the capital structure and the uncertainty in insurance market.

• Firm Size *SIZE*: the costly external equity suggests that it is more difficult for

[10]Net Premiums Earned record premium income for the year, prorated for the portion of the policy that occurs during the year in question. This is not a perfect proxy for liabilities, particularly in the case where policies are written on an occurrence basis. Where the firm has written occurrence policies in the past, the premiums have already been earned, but the liability still exists in that claims may still occur in the future.

smaller firms to issue equity in times of increasing of aggregate uncertainty; therefore, smaller insurance firms should tend to keep a higher equity-liability ratio. Warner (1977), Ang, Chua, and McConnel (1982) and Titman and Wessels (1988) provide evidence for non-financial firms that capital structure is related to firm size. One explanation for this is that transaction costs are decreasing in the size of the firm. Smith (1977) finds that small firms incur substantially more costs to issue equity than large firms. [11] In the regression, the natural logarithm of total admitted assets is used as a proxy for firm size. The predicted sign in the regression is positive.

• Past Profitability *PROFIT*: A positive cost of issuing equity, or a positive cost of distributing cash to shareholders implies a negative relationship between the capital structure and past profitability. This is because this positive cost of equity implies that the internally generated funds are low-cost source of equity capital for the insurance firm. The sample average of the profit/surplus ratios from 1999 to 2004 is used as a proxy of firm's past profitability.

Although there are 852 firms in our sample, a number of these firms are part of the same insurance group. Our sample contains 345 unique insurance groups. It is reasonable to assume that there will be some correlation of errors within each group. We correct for this by using the Huber-White sandwich estimator, which provides

[11] The transaction costs of issuing securities are defined as flotation costs and costs encountered in trying to secure the highest price for the firm's securities. Smith (1977) identified flotation costs as: (1) compensation paid to investment bankers, (2) legal fees, (3) accounting fees, (4) engineering fees, (5) trustee's fee, (6) listing fees, (7) printing and engraving fees, (9) federal revenue stamps, and (10) state taxes. Smith went on to provide evidence which showed that firms enjoy economies of scale when issuing securities.

robust standard errors

3.3.3 Results

The results are reported in Table 3.2.

DEPENDENT VARIABLE: NPE/E				
$R^2 = 0.3218$	Coefficient	Robust Std. Error	t	P>$\|t\|$
SD	-0.708*	0.067	-10.55	0.000
SIZE	0.680*	0.023	2.98	0.003
PROFIT	-0.638*	0.205	-3.12	0.002
Intercept	-3.214*	0.489	-6.57	0.000

Table 3.2: Results

The estimated elasticity of leverage with respect to uncertainty (SD) is -.708, which is both statistically and economically significant. This result confirms our hypothesis that leverage is indeed decreasing in the variance of firms' loss ratio. Firms faced with more uncertainty do choose capital structures which use less leverage.

The coefficient on past profit ($PROFIT$) is -.638, which is also statistically and economically significant. This confirms our second hypothesis, which is that firms that have greater access to internal capital (in this case, due to recent profitability) tend to use less leverage than do firms with a lesser supply of internal capital. This supports the notion that there is a cost advantage to internal equity, which is at the heart of previous studies of the behavior of insurance markets.

Finally, the coefficient on our variable controlling for $SIZE$ is 0.680. Larger firms use greater leverage than smaller firms, as predicted.

3.4 Conclusion

This paper explores the capital structure of insurers. The focus is on the impact of aggregate uncertainty, or dependence among risks, since this is the source of an insurer's incentive to issue equity. Insurance firms respond to the shocks of increased risks by taking all or some of the following actions: placing limits on the number or coverage of contracts that they offer; raising premium for the policies that they issue; and raising more equity. We analyze the equilibrium mixture of these responses in a competitive insurance market, and find that the impact of increasing uncertainty on the equity decision depends on the nature of aggregate uncertainty. Where this uncertainty is in the size of losses, equity increases with uncertainty; where the risk dependence is in the events of losses, equity first increases then decreases with uncertainty, providing that individuals are not too risk averse. The latter result follows from a tradeoff between two effects, which we label the input effect of uncertainty, and the output effect. In both cases, however, the ratio of equity to insurance revenue increases. We extend the model to look at the effect of a cost difference between internal equity (less costly) and external equity (more costly). This extension leads to the hypothesis that firms with greater internal equity will tend to use less leverage.

We test both hypotheses directly on a sample of 852 U.S. property and casualty stock insurers over a sample period from 1999-2004. We find support for both of our hypotheses. Firms that have higher variance in their loss ratio, our proxy for uncertainty, use significantly less leverage, supporting our theory that uncertainty and leverage are negatively correlated. Firms that have been recently profitable, implying greater internal capital, use significantly less leverage. This supports the theory that there is a cost advantage to internal over external equity, which is at the core of recent theories of insurance market dynamics.

3.5 Bibliography

[1] Ang, J., Chu J., and J. McConnell, 1982, "The Administrative Costs of Corporate Bankruptcy: A Note", *Journal of Finance* 37, 219-226.

[2] Arrow, K.J., 1971, "The Role of Securities in the Allocation of Risk-Bearing", *Essays in the Theory of Risk-Bearing,* Chicago, Markham.

[3] *Best's Aggregates and Averages,* annual 1980-1992.

[4] Bradley, M., Jarrell, G.A., and E. Han Kim, 1984, "On the Existence of an Optimal Capital Structure: Theory and Evidence", *Journal of Finance* 39, 857-78.

[5] Cummins, J.D., and S. Harrington, 1987, "The Relationship between Risk and Return: Evidence for Property-Liability Insurance", *Journal of Risk and Insurance,* 54.

[6] Doherty, Niel A., and Harris Schlesinger, 1990, "Rational Insurance Purchasing: Consideration of Contract Nonperformance", *The Quarterly Journal of Economics,* 243-253.

[7] Fischer, E., Heinkel, R. and J. Zechner (1989), "Dynamic Capital Structure Choice: Theory and Tests", *Journal of Finance* 44, 19-40.

[8] Grøn, A., 1994, "Capacity constraints and cycles in property-casualty insurance markets", *Rand Journal of Economics* 25, 110-127.

[9] Jensen, Michael C., and William H. Meckling, 1976, "Theory of the Firm: Agency costs, Managerial Behavior and Ownership Structure", *Journal of Financial Economics* 4, 305-360.

[10] Marshall, John M., 1976, "Insurance Theory: Reserves Versus Mutuality", *Economic Inquiry*, 476-492.

[11] Myers, Stewart, 1984, "The Capital Structure Puzzle", *Journal of Finance* 39, 575-92.

[12] Myers, Stewart, and N. Majluf, 1984, "Corporate Financing and Investment Decisions When firms have information investors do not have", *Journal of Financial Economics* 13, 187-221.

[13] OneSource, "U.S. Insurance and Reinsurance Transactions", 1-800-554-5501.

[14] Rajan, R.G. and L. Zingales, 1994, "What do we know about capital structure? Some evidence from international data" (Graduate School of Business, University of Chicago).

[15] Shyam-Sunder, Lakshmi, and Stewart C. Myers, 1994, "Testing Static Trade-off Against Pecking Order Models of Capital Structure", NBER working paper, No. 4722.

[16] Smith, C., Jr, 1977, "Alternative Methods for Raising Capital: Rights Versus Underwritten Offerings" *Journal of Financial Economics* 5, 273-307.

[17] Titman, S., 1984, "The Effects of Capital Structure on a Firm's Liquidation Decision", *Journal of Financial Economics* 13 (March 1984), 137-51.

[18] Titman, Sheridan and R. Wessels, 1988, "The Determinants of Capital Structure Choice", *The Journal of Finance* 43, 1-19.

[19] Warner, J., 1977, "Bankruptcy Costs: Some Evidence", *Journal of Finance* 32, 337-347.

[20] Winter, Ralph A. 1988, "The Liability Crisis and the Dynamics of Competitive Insurance Markets", *Yale Journal on Regulation* 5, 455-550.

[21] Winter, Ralph A. 1991, "Solvency Regulation and the Insurance Cycle", *Economic Inquiry.*

[22] Winter, Ralph A. 1994, "Dynamics of Competitive Insurance Markets", *Journal of Financial Intermediation.*

Chapter 4

Contracting With Agents of Heterogeneous Risk Aversion

4.1 Introduction

4.1.1 Overview

Firms' shareholders hire managers to look after their interests. Managers' tasks can be crudely divided into two categories. The first task is *project selection*, where managers decide what lines of business the firm ought to pursue, and what investments ought to be made. The second task is *effort exertion*, where managers can improve the distribution of eventual project outcomes by working harder. It is impossible for shareholders to know whether managers' actions were optimal from their perspective on either of these tasks. With respect to project selection, the manager has an informational advantage that comes about either because the manager was specifically hired for his expertise in this area, or because his position affords him the opportunity develop a better knowledge of the firm's opportunities than anyone else. It is impossible for shareholders to know whether the manager's project selection decision was the "right" one. With respect to effort exertion, it is assumed that shareholders simply cannot monitor the manager's level of effort.

Dealing with these problems is the standard purview of the principal-agent literature. The solution, particularly with respect to motivating the proper effort level on the part of the manager, is to make part of the manager's pay package depend on firm performance. The advantage of incentive pay is that it aligns shareholder and managerial interests; the disadvantage is that it involves a deviation from optimal risk sharing. Well diversified shareholders are presumed to be risk neutral with respect to the firm's idiosyncratic risk. Risk averse managers are unable to diversify their exposure to firm risk, meaning that they place less value on risky pay than it is expected to cost the firm's shareholders to provide it. The standard approach to determining the optimal contract is trade off the costs and benefits of incentive pay, choosing the level of power (i.e. amount of incentive pay provided) at which the marginal costs equal the marginal benefits.

I motivate this paper with a number of observations. The first is that managers differ in their degree of risk aversion, and that a manager's risk aversion is not observable. This complicates the problem of choosing the correct tradeoff between inducing managerial effort and deviating from optimal risk sharing.

Second, higher managerial risk aversion is costly in two ways. First, higher risk aversion means that the manager puts a lower value on risky pay. This implies that the cost of motivating effort exertion is increasing in managerial risk aversion. However, managerial risk aversion is also costly in terms of motivating correct *project selection.* When selecting projects, managers have an incentive to choose those that best fit their own interests, as opposed to those of firm shareholders. This becomes important when projects differ in dimensions such as the degree of risk they impose. The greater difference in risk preferences between risk averse managers and risk neutral shareholders, the greater will be the distortion imposed by managers selecting projects according to their own interests. As such, shareholders prefer managers with lower risk aversion for two reasons: it is less costly to motivate effort exertion, and these managers' project selection decisions will more closely match shareholders' preferred outcomes.

A third observation is that the market for managerial labour, like any labour market, is a competitive one. Firms compete with one another for the services of preferred managers, and managers will choose to work for the firm that makes them the offer they prefer.

These three observations taken together imply the following. From the second observation, it is clear that firms prefer lower risk aversion managers. From the third observation, they must compete against other firms for the services of lower risk aversion managers. And from the first observation, such competition is difficult, since a manager's risk aversion is his own private information. Firms must therefore develop contracts which serve as screening devices, designed so that they will attract low risk aversion managers. Since all managers prefer more pay to less, firms cannot compete for low risk aversion managers simply by raising wages. If they wish to separate relatively desirable low risk aversion managers from relatively undesirable high risk aversion managers, they must compete in a manner that exploits the differences between types. Since lower risk aversion managers put greater value on risky pay than high risk aversion managers, firms have an incentive to offer high-powered contracts as a screening device. Such contracts appeal to the targeted low risk aversion managers, but not to high risk aversion managers.[1]

This paper explores the impact that this selection effect has on the design of managerial contracts. I develop a model where firms must compete against one another in the managerial labour market to attract managers who are responsible for both project selection and effort exertion. In this setting, incentive contracts perform two functions. The first is to serve the traditional role of motivating the correct effort choice. The second is to act as a screening mechanism, helping firms compete for

[1]Screening models are more traditionally thought of in the context of insurance contracts, where an insurer sets out a menu of contracts to offer to customers who walk in the door. The practice of hiring a CEO is clearly a much more selective one, and firms put enormous effort into learning as much as possible about prospective candidates. However, for any executive position the firm will identify a number of candidates, whose risk aversion will likely remain difficult to discern ex ante. Similarly, candidates for top positions generally appeal to more than one prospective employer. As such, a competitive screening model is a reasonable approach to modelling the problem.

the services of a lower risk aversion manager whose preferences lead to better project selection.

4.1.2 Literature review

Principal-agent problems have long been studied in the literature. Papers by Jensen and Meckling, Ross, Holmstrom, and Holmstrom and Milgrom are well-known early examples which highlight the difficulties of contracting between a principal and his agent when the agent's actions are not easily observed. The traditional principal-agent model developed in most of these (and later) papers is one where an agent, by exerting costly effort, is able to improve the expected outcome of a principal. This is often described as the agent exerting "productive effort". A less common, but very interesting, related form of model is based instead on the agent exerting effort in order to evaluate a number of potential projects. The principal must motivate the agent to expend effort to examine several opportunities, and then implement one of them, with the principal unable to observe the agent's selection process or decision criteria. Lambert's (1986) model of "evaluation effort" is one of the earlier models that capture this idea, which has also been modeled recently by Core and Qian (2002). Papers examining the compensation of investment fund managers often use a similar structure.

The early research into principal-agent problems has been taken up with enthusiasm recently as efforts to explain the nature of executive compensation have grown almost as quickly as the compensation itself. Murphy (1999), Hall and Murphy (2001) and Core, Guay and Larcker (2002) provide excellent surveys of the voluminous literature that has exploded around the general question of why executives are being compensated as they are, and whether the compensation they receive is consistent with optimal contracting.

An important research question is to ask how managers behave when provided with risky compensation. Numerous papers have argued that this type of pay encourages risk taking. Taken to the extreme, consider risk neutral managers who are provided with a call option on an asset whose underlying volatility they control. Since the risk neutral value of a call option is increasing in underlying volatility, the managers would choose as risky an investment strategy as they possibly could.

This solution isn't very satisfying, particularly because most of the many managers compensated with share options do not seem to be trying to drive their firms' volatility to unprecedented and dizzying levels.[2] Several recent papers have introduced the notion that an executive's risk aversion should be considered when analyzing how risky pay will affect his project or investment selection. Since the executive has a great deal of wealth tied up in firm-specific securities, in addition to his human capital being highly correlated with firm performance, it's not unreasonable to believe that a manager will be very concerned about his firm's idiosyncratic risk. Carpenter

[2]The Skillings and Enrons of this world remain more the exception than the rule, although some might wish to debate this.

(2000) finds that a mutual fund manager who is risk averse with respect to investment performance will, in some circumstances, actually behave *more* conservatively if given more options. Recently, both Ross (2004) and Lewellen (2003) explore how risk aversion filters the effect of risky compensation on managerial decision making. These and other papers make the case that to understand the effect of *any* compensation package, a manager's private preferences toward risk are a crucial element that must be considered.

Within the context of the executive compensation literature, relatively little has been done with regard to considering how differences in agents' risk aversion affects the design of pay packages. An exception is Jullien et al (2000), which considers executive compensation as one application of their model describing how a risk neutral principal ought to contract with a number of risk averse agents having heterogeneous, private levels of risk aversion. Unlike this paper, their model assumes that the firm takes on numerous principal-agent relationships, as opposed to contracting with only one agent. Serfes (2005) considers a matching game between risk neutral principals and agents of differing risk aversion in a labour market setting. However, in this model agents exert only productive effort, and do not make project selection decisions. Wright (2004) presents a model similar to that in this paper, in terms of a setting featuring two types of agents with firms competing for their services. Again, this model takes firm risk as exogenous, and agents make no project selection decisions. Both the Serfes and Wright papers predict that agents of lower risk aversion will contract with firms that are *a priori* riskier.

Understanding risk aversion, and differences in risk aversion between agents, is important when considering results from empirical compensation studies. The Jensen and Murphy (1990) result asserting that executive pay was "lower" than it ought to have been is well known. But interpreting the degree to which pay changes relative to changes in firm wealth must be taken in the context of executive risk aversion. Haubrich (1994) attempts to measure this for a single type of agent, while this paper emphasizes that differences in risk aversion and the associated constraints on contracting should be considered as well.

As well, there is some question as to why the relationship between the degree to which pay is risky and the level of firm-specific risk is so tenuous. A standard argument in the literature is that if executives are risk averse, then they should be asked to bear less risky compensation as firm risk increases. Prendergast (2002) surveys the empirical literature on this question and finds the evidence is decidedly mixed (three studies find the predicted negative relationship, three a positive relationship, and six no statistically significant relationship). His paper argues that the ambiguity of the contract setting is a contributing factor. I suggest that differences in risk aversion may also play a role. For example, risky high technology firms, faced with hiring a manager to select projects in an unstructured environment, have a strong incentive to bid for low risk aversion agents. When contracting with low risk aversion agents, these firms can take advantage of the agent's risk tolerance by offering high powered incentive contracts, despite the relatively high firm risk. Less risky firms contracting with higher risk aversion agents would then not necessarily offer higher powered contracts, despite these firms' lower return volatility. Thus the relationship between firm risk and power of incentive contracts would not necessarily be negative, as standard theory predicts.

4.2 Model: Single Firm, Single Agent

Consider a firm with the opportunity to hire a manager to select a one-period investment project, and then exert effort to implement the project. The manager will choose between a safe project and a risky project, the safe project returning a value of v_0, while the risky project's terminal value will be either v_g or v_b. The prior probability of obtaining the high or low outcomes is $1/2$ for each. So that one project does not dominate the other *ex ante*, $v_b < v_0 < v_g$. For ease of exposition, the risk free interest rate is assumed to be zero.

The firm's shareholders are well diversified, and therefore risk neutral with respect to the idiosyncratic risk posed by the uncertain project. Prospective managers' prospects for diversification are much less, and they are therefore assumed to be risk averse with respect to an employment contract that calls for their wage to have a stochastic component.

Once hired, the chosen manager receives a signal $\tilde{r}$, which is the probability that the risky project will return v_g. $\tilde{r}$ is distributed uniformly over $[0, 1]$. The signal is the manager's private information, and he is unable to communicate this signal credibly to firm shareholders.

With the updated signal, the manager chooses between the safe project and the risky project. If he chooses the safe project, the firm's terminal asset value is v_0 with certainty. The manager does not need to exert any effort in this case. However, if the manager chooses the risky project, he has the opportunity to exert effort, at a personal cost of c dollars, to improve the probability of success. If the manager exerts effort, the probability of realizing the high asset value v_g is r (the realization of the manager's signal), while the low value v_b will occur with probability $1 - r$. If the manager does not exert effort, the project will certainly fail, and the terminal value of the project is v_b with probability 1.

The manager's effort is not observable to firm shareholders. To provide the manager with an incentive to exert effort, the firm must offer the manager an incentive contract; that is, the manager's payoff must depend on the firm's terminal asset value. A contract $\mathbf{S}$ therefore takes the form $\mathbf{S} = [s_0, s_b, s_g]$, defining the payment that the manager receives for each possible outcome in firm asset value.[3]

[3]When productive effort is removed from the model and effort is shifted to the first stage of the game, the model collapses to that of Lambert (1986). Lambert's solution differs markedly from the joint selection-production model developed here.

4.2.1 First best solution

If the shareholders could observe the manager's signal, and determine whether the manager working on the good project exerted effort, the first best result would be possible. At the project selection stage of the game, the investment policy taken by the manager can be described in terms of a cutoff point p. When the signal is below this point, the prospect of the risky project succeeding is too low, and the shareholders would prefer that the manager pursue the safe project. When the signal is above p, the probability of success is sufficiently high, and the shareholders would prefer that the manager pursue the risky project.

In the full information case, the firm pays the manager a fixed wage, $\bar{w}$, equal to the manager's reservation wage. Should the signal indicate that the risky project is worth pursuing, the firm pays the manager a bonus of c to compensate the manager for the effort required to make the good realization possible.[4] The first best rule is to select the risky project if the realization of $\tilde{r}$ is such that

$$rv_g + (1-r)v_b - c \geq v_0$$

Taking the first best cutoff point p^*_{FB} as the value of r which makes this an equality and rearranging yields

$$p^*_{FB} = \frac{v_0 - v_b + c}{v_g - v_b} \tag{4.1}$$

[4]Since both the signal and effort are observable in this case, there is no principal-agent conflict. The manager pursues exactly the investment policy that the shareholders desire, and effort is verifiable and therefore contractible.

The strategy of making the cutoff point, and taking the risky project if and only if the realization of $\tilde{r}$ satisfies $r \geq p^*_{FB}$, maximizes expected asset value net of the expected cost of effort.

4.2.2 Hidden information: the decisions of risk averse managers

When the manager's signal and effort level are not observable, the payment to the manager can depend only on the realization of firm asset values. The contract must not only provide the manager with an incentive to exert effort in the appropriate circumstances (i.e. when he selects the risky project), it must also elicit the correct project selection decision at the first stage of the game. While this would be simple if the manager were risk neutral, in reality managers are risk averse. The problem becomes more complex when this risk aversion is taken into account.

The manager is assumed to have negative exponential (CARA) utility of the form

$$U(w) = -e^{-\gamma w} \tag{4.2}$$

where γ is the coefficient of absolute risk aversion. This functional form has the property that the manager's decisions will not change as his level of wealth changes, which provides convenient tractability when working to solve both the firm's and the manager's respective maximization problems (see Holmstrom and Milgrom 1987).

The manager is assumed to have an outside option which provides a certain payment of $\bar{w}$. The manager will therefore not accept any contract providing ex ante expected utility of less than $U(\bar{w}) = \bar{U}$. However, once the manager accepts the contract at the start of the game, the outside option disappears.[5]

When offered a contract $\mathbf{S} = [s_0, s_b, s_g]$, the manager will choose the investment cutoff p to maximize his own utility. That point is where the manager is indifferent between pursuing the safe project and exerting effort on the risky project:

$$U(s_0) = (1-p)U(s_b - c) + pU(s_g - c) \tag{4.3}$$

This implies that on a given contract, the manager's chosen cutoff point p is given by

$$p = P(\mathbf{S}) = \frac{U(s_0) - U(s_b - c)}{U(s_g - c) - U(s_b - c)} \tag{4.4}$$

4.2.3 The optimal contract

The following expressions are useful in the derivation of the optimal contract. Recall that the updated signal of risky project's probability of success, $\tilde{r}$, is distributed uniformly over $[0, 1]$. However, the *ex ante* probability of the ultimate outcome being the successful risky project depends not only on the realization of the updated signal, but also on whether or not the signal is greater than the cutoff point p. Since contracting decisions are made prior to the transmission of the updated signal, it is

[5]This rules out a strategy where the manager waits to observe the signal and quits if it is not to his liking.

helpful to derive *ex ante* probabilities for each of the three outcomes (safe project v_0, successful risky project v_g, and unsuccessful risky project v_b) given a cutoff point p.

Since the signal $\tilde{r}$ is distributed uniformly over $[0,1]$, and the safe project is pursued for any realization below p, the probability of pursuing the safe project is $F(p) = p$, where F is the uniform CDF of $\tilde{r}$. The other two outcomes first require that the signal be greater than p. The *ex ante* probability of realizing the good distribution conditional on the project's signal meeting the cutoff point is $\int_p^1 rf(r)dr = \frac{1}{2}(1-p^2)$. A similar argument can be made to show that the prior probability of realizing the bad return is $\frac{1}{2}(1-p)^2$.

Let $V(p)$ be the expected value of the firm's assets under investment policy p, and $C(\mathbf{S};p)$ be the expected wage cost of offering contract $\mathbf{S}$ when the manager's investment cutoff policy is p. The firm's objective is to maximize expected profit, $\pi(\mathbf{S};p)$, where

$$\pi(\mathbf{S};p) = V(p) - C(\mathbf{S};p) \tag{4.5}$$

and

$$\begin{aligned} V(p) &= pv_0 + \frac{1}{2}(1-p^2)v_g + \frac{1}{2}(1-p)^2 v_b && (4.6) \\ C(\mathbf{S};p) &= ps_0 + \frac{1}{2}(1-p^2)s_g + \frac{1}{2}(1-p)s_b && (4.7) \end{aligned}$$

where the probabilities associated with each outcome are the *ex ante* probabilities of the outcome occurring given the investment policy p.

The maximization problem is subject to the following constraints:

$$pU(s_0) + \frac{1}{2}(1-p^2)U(s_g - c) + \frac{1}{2}(1-p)^2 U(s_b - c) \geq \bar{U} \quad (4.8)$$

$$(1-p)U(s_b - c) + pU(s_g - c) \geq U(s_b) \quad (4.9)$$

$$(1-p)U(s_b - c) + pU(s_g - c) = U(s_0) \quad (4.10)$$

Equation 4.8 is the manager's incentive compatibility constraint, requiring that *ex ante* the manager's expected utility from pursuing the investment policy p under the contract meet his level of reservation utility. Equations 4.9 and 4.10 are the individual rationality constraints which govern behaviour at the second stage of the game, after the manager receives the signal. Equation 4.9 requires that the manager prefer to work under the contract than simply select the risky project and exert no effort. Equation (4.10) requires that the manager not prefer to exert effort under the risky project rather than avoid effort by selecting the safe project when the signal is equal to the cutoff point.[6]

4.2.4 Solution properties

Proposition 4.1 *In the single agent, single firm case, 4.8 binds in equilibrium.*

Proof: Equation 4.8 is the incentive compatibility constraint, and holds that the manager's expected utility upon taking the contract must meet his level of reservation

[6] There are a continuum of incentive compatiblity constraints with respect to the payout from the safe project, one for each realization of the signal $\tilde{r}$. In equilibrium, all are redundant except for 4.10, where $r = p$. 4.10 is a rearrangment of the condition from 4.4 requiring that the manager's chosen cutoff point p satisfy $P(\mathbf{S})$.

utility. Under negative exponential utility, a utility function of the Constant Absolute Risk Aversion (CARA) class, the decisions the manager makes with respect to investment and effort decisions depend on the *relative differences* between s_0, s_b and s_g, not on their absolute levels.

Consider any contract $\mathbf{S}$,motivating cutoff point $p = P(\mathbf{S})$, for which the incentive compatibility constraint 4.8 does not bind. There must always exist a contract $\acute{\mathbf{S}} = \mathbf{S} - \boldsymbol{\epsilon}$, under which ϵ is subtracted from the payment in every state made under the original contract. ϵ is sufficiently small such that 4.8 remains satisfied. Taking 4.4 and substituting $U(w) = -e^{-\gamma w}$, both contracts motivate the same decision $p = P(\mathbf{S}) = P(\acute{\mathbf{S}})$:

$$\begin{aligned} P(\acute{\mathbf{S}}) &= \frac{-e^{-\gamma(s_0-\epsilon)} - \left(-e^{-\gamma(s_b-c-\epsilon)}\right)}{\left(-e^{-\gamma(s_b-c-\epsilon)}\right) - \left(-e^{-\gamma(s_b-c-\epsilon)}\right)} \\ &= \frac{\left(-e^{-\gamma(s_0)} - \left(-e^{-\gamma(s_b-c)}\right)\right) e^{-\gamma(-\epsilon)}}{\left(\left(-e^{-\gamma(s_b-c)}\right) - \left(-e^{-\gamma(s_b-c)}\right)\right) e^{-\gamma(-\epsilon)}} \\ &= P(\acute{\mathbf{S}}) \end{aligned}$$

Then it must be that

$$\pi(\mathbf{S};p) < \pi(\acute{\mathbf{S}};p)$$

as the expected asset value is the same under each contract, while the expected wage cost is lower under $\acute{\mathbf{S}}$. $\mathbf{S}$ cannot be optimal. *QED*

Proposition 4.2 *At an optimal contract $\hat{S}$, $\hat{p} = P(\hat{S}) \geq p^*_{FB}$.*

Proof: Define

$$\begin{aligned}\Psi(\mathbf{S};p) &= pU(s_0)+\frac{1}{2}(1-p^2)U(s_g-c)+\frac{1}{2}(1-p)^2U(s_b-c)\\ \text{for } \mathbf{S} &= (s_0,s_g,s_b)\end{aligned}$$

Consider a contract $\acute{\mathbf{S}}$ such that $P(\acute{\mathbf{S}})=\acute{p}<p^*_{FB}$. Consider a deviation from $\acute{\mathbf{S}}$ given by $\breve{\mathbf{S}}=(\breve{s}_0,\breve{s}_g,\breve{s}_b)=(\acute{s}_0+\frac{\frac{1}{2}(1-\acute{p}^2)}{\acute{p}}\epsilon,\acute{s}_g-\epsilon,\acute{s}_b)$. Then evaluate Ψ for $\breve{\mathbf{S}}$ under the old $\acute{p}$. This expression is greater than it was under $\acute{\mathbf{S}}$:

$$\Psi(\acute{\mathbf{S}};\acute{p})<\Psi(\breve{\mathbf{S}};\acute{p})$$

since $\acute{\mathbf{S}}$ is a mean preserving spread of $\breve{\mathbf{S}}$.[7] Because $P(\acute{\mathbf{S}})\neq P(\breve{\mathbf{S}})$, the agent's equilibrium investment policy under $\breve{\mathbf{S}}$ changes and is given by $\breve{p}=P(\breve{\mathbf{S}})$. Because $\breve{p}$ is utility maximizing under $\breve{\mathbf{S}}$

$$\Psi(\breve{\mathbf{S}};\acute{p})<\Psi(\breve{\mathbf{S}};\breve{p})$$

Define t such that

$$\Psi(\breve{\mathbf{S}};\acute{p})=\Psi(\breve{s}_0-t,\breve{s}_g-t,\breve{s}_b-t;\breve{p})$$

Then

$$\pi(\acute{\mathbf{S}};\acute{p})=\pi(\breve{\mathbf{S}};\acute{p})<\pi(\breve{\mathbf{S}}-\mathbf{t};\breve{p})$$

The first and second terms are equal since both are evaluated at $\acute{p}$ and the spread

[7] Note that while $\Psi(\mathbf{S};p)$ can be evaluated for any p, it is not necessarily the case that p satisfies $P(\mathbf{S})$. However, for $\Psi(\mathbf{S};p)$ to represent the true equilibrium expected utility from $\mathbf{S}$, p must satisfy $p=P(\mathbf{S})$.

from $\breve{\mathbf{S}}$ to $\acute{\mathbf{S}}$ is mean preserving.[8]

$$P(\breve{\mathbf{S}}) > P(\acute{\mathbf{S}})$$

because the individual rationality constraint (4.9) is satisfied for fewer realizations of $\tilde{r}$ under $\breve{\mathbf{S}}$ than under $\acute{\mathbf{S}}$. $P(\acute{\mathbf{S}}) < p^*_{FB}$ and $P(\acute{\mathbf{S}}) < P(\breve{\mathbf{S}})$, therefore $V(\acute{p}) < V(\breve{p})$. For ϵ sufficiently small, all surplus from the move from ($\breve{\mathbf{S}}$,p) to ($\breve{\mathbf{S}} - \mathbf{t}$,$\hat{p}$) accrues to the firm. Thus given any contract $P(\acute{\mathbf{S}}) < p^*_{FB}$ a Pareto dominant contract exists, and $\acute{\mathbf{S}}$ cannot be optimal. *QED*

Proposition 4.3 *In the single agent, single firm case, 4.9 binds in equilibrium.*

Proof: Consider a contract $\acute{\mathbf{S}} = (\acute{s}_0, \acute{s}_g, \acute{s}_b)$ for which 4.9 does not bind. Let $\acute{p} = P(\acute{\mathbf{S}})$. For small ϵ, define $\breve{\mathbf{S}}$ such that

$$\begin{aligned}
\breve{s}_0 &= \acute{s}_0 - \frac{\frac{1}{2}(1-\acute{p})^2}{\acute{p}}\epsilon \\
\breve{s}_g &= \acute{s}_g \\
\breve{s}_b &= \acute{s}_b + \epsilon
\end{aligned}$$

so $\acute{\mathbf{S}}$ is a mean preserving spread of $\breve{\mathbf{S}}$ given $\acute{p}$. Then

$$\pi(\acute{\mathbf{S}};\acute{p}) = \pi(\breve{\mathbf{S}};\acute{p})$$

but

$$\Psi(\acute{\mathbf{S}};\acute{p}) < \Psi(\breve{\mathbf{S}};\acute{p})$$

[8]Note that $\pi(\breve{\mathbf{S}}, \acute{p})$ cannot be an equilibrium level of profit because the expression is evaluated at $\acute{p}$ and $\acute{p} \neq P(\breve{\mathbf{S}})$.

because the risk averse agent prefers to avoid the lottery presented by the mean preserving spread in $\acute{\mathbf{S}}$. Because $\acute{p}$ is not a utility maximizing investment level for contract $\breve{\mathbf{S}}$, since $P(\breve{\mathbf{S}}) = \breve{p}$, we have

$$\Psi(\breve{\mathbf{S}}; \acute{p}) < \Psi(\breve{\mathbf{S}}; \breve{p})$$

Define $\mathbf{t}$ such that

$$\Psi(\breve{\mathbf{S}} - \mathbf{t}; \breve{p}) = \Psi(\breve{\mathbf{S}}; \acute{p})$$

and note that $\breve{p} = P(\breve{\mathbf{S}}) = P(\breve{\mathbf{S}} - \mathbf{t})$ under CARA utility. Finally,

$$\pi(\breve{\mathbf{S}} - \mathbf{t}; \breve{p}) > \pi(\breve{\mathbf{S}}; \acute{p}) = \pi(\acute{\mathbf{S}}; \acute{p})$$

because $p^*_{FB} < \breve{p} < \acute{p}$, and the downward shift of t in C allows the principal to capture the surplus. Therefore, when 4.9 does not bind for $\acute{\mathbf{S}}$, there exists a Pareto dominant contract that leaves the agent's utility unchanged and makes the principal strictly better off. *QED*

Corollary 4.1 At the optimum $s_0 = s_b$. That is, the wage paid for selecting the safe project is set equal to the wage paid when the risky project is selected and the bad return is realized.

Proof: Equation 4.9 binds and 4.10 is an equality. Both are identical on the left hand side, so $U(s_b) = U(s_0)$. By monotonicity of the utility function, $s_b = s_0$. *QED*

Since equations 4.8 through 4.10 bind, it is possible to determine the minimum cost contract to motivate a given investment policy p. Substitute $s_b = s_0$ into 4.4:

$$p = P(\mathbf{S}) = \frac{U(s_0) - U(s_0 - c)}{U(s_g - c) - U(s_0 - c)} \tag{4.11}$$

Define $g = s_g - s_0$, and since $U(w) = -e^{-\gamma w}$, one can factor out $e^{-\gamma s_0}$ and write

$$p = P(\mathbf{S}) = \frac{U(0) - U(-c)}{U(g - c) - U(-c)} \tag{4.12}$$

Solving this expression for g yields

$$g = G(p) = \frac{c\gamma - Ln\left[\frac{1 - e^{c\gamma}(1-p)}{p}\right]}{\gamma} \tag{4.13}$$

Define $w = s_0$ as the fixed wage for a given contract $\mathbf{S}$. For a given g, the firm chooses w so as to strictly satisfy the incentive compatibility constraint(4.8). The dimension of the contract space is reduced to a fixed wage w, and a bonus g paid when the good state is realized. I abuse notation by retaining $\mathbf{S}$ to denote (now) two dimensional contracts. Efficient contracts therefore take the form $\mathbf{S} = (w, g)$, where g is the bonus that motivates the agent to pursue investment cutoff policy p, and w is the wage required to meet the agent's reservation utility strictly.

The expected wage cost of motivating an investment cutoff point p is

$$C(\mathbf{S}; p) = w + (1 - p^2)g$$

where $g = G(p)$. It is therefore possible to calculate the expected wage cost, $C(\mathbf{S}; p)$, of motivating a given agent to pursue any cutoff point p, as well as the expected asset value, $V(p)$, from the cutoff point, p. The firm then chooses the contract $\hat{\mathbf{S}}$ which motivates the cutoff point $\hat{p}$, at which point the marginal expected wage cost of a change in p is equal to the marginal increase in expected firm value, $\frac{\partial C}{\partial p}_{p=\hat{p}} = \frac{\partial V}{\partial p}_{p=\hat{p}}$.

QED

Lemma 4.1 *The maximum expected utility (measured in terms of a certainty equivalent payment) an agent can derive from any contract is decreasing in the agent's risk aversion.*

Proof: Let $\Psi_i(\mathbf{S};p)$ be the expected utility that an agent of risk aversion γ_i derives from contract $\mathbf{S}$ evaluated at p.

$$\Psi_i(\mathbf{S};p) = pU_i(w) + \frac{1}{2}(1-p^2)U_i(w+g-c) + \frac{1}{2}(1-p)^2U_i(w-c)$$

Let $CE_i(\Psi_i)$ be the certainty equivalent of any expected utility Ψ_i for an agent of risk aversion γ_i:

$$CE_i(\Psi_i) = U_i^{-1}(\Psi_i)$$

Consider an agent L having risk aversion γ_L, and an agent H having higher risk aversion $\gamma_H = \gamma_L + \epsilon$. Consider the expected utility either type derives from contract $\mathbf{S}$ by pursuing investment policy $p_H = P_H(\mathbf{S})$, where p_H is the *high type's* utility maximizing investment cutoff point under the contract. Using the fact that the agents' utility functions are of the form $U_i = -e^{-\gamma_i}$:

$$\Psi_i(\mathbf{S};p) = pe^{-\gamma_i(w)} + \frac{1}{2}(1-p^2)e^{-\gamma_i(w+g-c)} + \frac{1}{2}(1-p)^2e^{-\gamma_i(w-c)}$$

This is equivalent to the utility from a fixed payment w and a lottery paying $-c$ with probability $\frac{1}{2}(1-p)^2$, 0 with probability p and $(g-c)$ with probability $\frac{1}{2}(1-p^2)$. Each agent places the same value on the certain payment w, and the difference in certainty equivalent utility derived under the contract depends on the value each places on the lottery. Because H is more risk averse than L, $U_H(x)$ is a concave transformation of $U_L(x)$. Then it must be the case that L is willing to pay more than H for any lottery. Therefore

$$CE_L(\Psi_L(\mathbf{S}; p_H)) > CE_H(\Psi_H(\mathbf{S}; p_H))$$

To complete the proof, observe that while type H pursues his expected utility maximizing investment policy $p_H = P_H(\mathbf{S})$, the policy is not utility maximizing for L, whose utililty maximizing policy p_L satisfies $p_L = P_L(\mathbf{S})$ Therefore

$$\Psi_L(\mathbf{S}; p_L) > \Psi_L(\mathbf{S}; p_H)$$

It follows that

$$CE_L(\Psi_L(\mathbf{S}; p_L) > CE_L(\Psi_L(\mathbf{S}; p_H)) > CE_H(\Psi_H(\mathbf{S}; p_H))$$

which proves the lemma.[9] *QED*

Proposition 4.4 *The expected wage cost of motivating any investment policy $p < 1$, is increasing in γ, the manager's risk aversion.*

Proof: Consider an agent L having risk aversion γ_L, and an agent H having higher risk aversion $\gamma_H = \gamma_L + \epsilon$. Both demand the same certainty-equivalent wage,

[9] If p were fixed, this lemma would be simply a representation of one of the outcomes of increasing risk aversion described in Rothchild Stiglitz (1970). It is the endogeneity of p that makes this result non-trivial.

$\bar{w}$. Let $\acute{\mathbf{S}}_L = (\acute{w}_L, \acute{g}_L)$ be the least cost contract that motivates type L to pursue investment policy $\acute{p}$, and let $\acute{\mathbf{S}}_H = (\acute{w}_H, \acute{g}_H)$ be a contract motivates type H to pursue $\acute{p}$. To prove the proposition, consider a contradiction: $C(\acute{\mathbf{S}}_H; \acute{p}) = C(\acute{\mathbf{S}}_L; \acute{p})$. From 4.13, $\frac{\partial g}{\partial \gamma} > 0$, so $\acute{g}_H > \acute{g}_L$. Let

$$\Delta\acute{g} = \acute{g}_H - \acute{g}_L \tag{4.14}$$

and

$$\Delta\acute{w} = \acute{w}_L - \acute{w}_H \tag{4.15}$$

Define the difference in cost under the two contracts as

$$\Delta C = C(\acute{\mathbf{S}}_H; \acute{p}) - C(\acute{\mathbf{S}}_L; \acute{p}) = 0 \tag{4.16}$$

Since $\Delta C = 0$ if there is no difference in cost, then

$$\Delta\acute{w} = \frac{1}{2}(1 - \acute{p}^2)\Delta\acute{g} \tag{4.17}$$

$\acute{\mathbf{S}}_H$ is therefore a mean preserving spread of $\acute{\mathbf{S}}_L$. Because the agents are risk averse,

$$\Psi_H(\acute{\mathbf{S}}_L; \acute{p}) > \Psi_H(\acute{\mathbf{S}}_H; \acute{p}) \tag{4.18}$$

Consider the utility of agent H under $\acute{\mathbf{S}}_L$. Let $\breve{p} = P_H(\acute{\mathbf{S}}_L)$ be the utility maximizing investment cutoff for agent H under $\acute{\mathbf{S}}_L$. Because $\acute{p}$ is not utility maximizing for type H under the contract

$$\Psi_H(\acute{\mathbf{S}}_L; \acute{p}) < \Psi_H(\acute{\mathbf{S}}_L; \breve{p}) \tag{4.19}$$

From the previous lemma and 4.19

$$CE_H(\Psi_H(\acute{\mathbf{S}}_L; \acute{p})) < CE_L(\Psi_L(\acute{\mathbf{S}}_L; \acute{p})) \tag{4.20}$$

Because $\acute{\mathbf{S}}_L$ is the least cost contract to motivate L to pursue $\acute{p}$, the incentive compatibility constraint must bind, therefore

$$CE_L(\Psi_L(\acute{\mathbf{S}}_L; \acute{p})) = \bar{w} \tag{4.21}$$

Then

$$CE_H(\Psi_H(\acute{\mathbf{S}}_L; \acute{p})) < \bar{w} \tag{4.22}$$

Because $\acute{\mathbf{S}}_H$ is a mean preserving spread of $\acute{\mathbf{S}}_L$

$$\Psi_H(\acute{\mathbf{S}}_H; \acute{p}) < \Psi_H(\acute{\mathbf{S}}_L; \acute{p}) \tag{4.23}$$

then

$$CE_H(\Psi_H(\acute{\mathbf{S}}_H; \acute{p})) < CE_H(\Psi_H(\acute{\mathbf{S}}_L; \acute{p})) < \bar{w} \tag{4.24}$$

Since

$$\Psi_H(\acute{\mathbf{S}}_H; p) < U_H(\bar{w}) \tag{4.25}$$

$\acute{\mathbf{S}}_H$ is not an equilibrium contract. The minimum cost contract to motivate H to pursue $\acute{p}$ is some contract $\hat{\mathbf{S}}_H = (\hat{w}_H = \acute{w}_H + t, \hat{g}_H = \acute{g}_H)$ with t defined so

$$\Psi_H(\hat{\mathbf{S}}_H; p) = U_H(\bar{w}) \tag{4.26}$$

Since t is a fixed wage payment, and $\mathbf{S}_H$ is a mean preserving spread of $\mathbf{S}_L$

$$C(\hat{\mathbf{S}}_H; \acute{p}) > C(\mathbf{S}_H; \acute{p}) = C(\mathbf{S}_L; \acute{p}) \tag{4.27}$$

Since this is true for any ϵ increase in risk aversion this proves the proposition, so

$$\frac{\partial C(\mathbf{S}; P(\mathbf{S}))}{\partial \gamma} > 0 \tag{4.28}$$

QED

Proposition 4.5 *The expected wage cost of any least cost contract S motivating p is decreasing in p.*

Proof: Consider a least cost contract $\acute{\mathbf{S}} = (\acute{w}, \acute{g})$ that motivates $p = P(\acute{\mathbf{S}})$, and a second least cost contract $\breve{\mathbf{S}} = (\breve{w}, \breve{g})$ that motivates $\breve{p} = P(\breve{\mathbf{S}})$, where $\breve{p} = \acute{p} - \epsilon$.[10] From 4.13, $\frac{\partial g}{\partial p} < 0$, so $\breve{g} > \acute{g}$. Let Δg be the difference in the expected value of the bonus payment between $\breve{S}$ and $\acute{S}$:

$$\Delta g = \frac{1}{2}(1 - \acute{p}^2)(\breve{g} - \acute{g}) + \frac{1}{2}[(1 - \breve{p}^2) - (1 - \acute{p}^2)]\breve{g} \qquad (4.29)$$

Let

$$\Delta w = \acute{w} - \breve{w} \qquad (4.30)$$

Taking contract $\breve{\mathbf{S}}$ is equivalent to taking contract $\acute{\mathbf{S}}$ (from which the agent derives reservation utility $U(\bar{w})$) and paying Δw for a lottery with expected payout Δg. Let

$$\Psi(\Delta g) = \frac{1}{2}(1 - \acute{p}^2)U(\breve{g} - \acute{g}) + \frac{1}{2}[(1 - \breve{p}^2) - (1 - \acute{p}^2))]U(\breve{g}) \qquad (4.31)$$

be the agent's expected utility from lottery. Because $\Psi(\acute{\mathbf{S}}; \acute{p}) = \Psi(\breve{\mathbf{S}}; \breve{p})$ and the agent has negative exponential utility

$$U(\Delta w) = \Psi(\Delta g) \qquad (4.32)$$

[10] Note that a lower p is a more aggressive investment policy, since the risky project will be chosen for more realizations of $\tilde{r}$.

Because the utility function is concave, one can apply Jensen's Inequality to the left hand side of 4.31 to see that

$$U(\Delta g) > \Psi(\Delta g) \tag{4.33}$$

Combining 4.32 and 4.33, it must be that $\Delta g > \Delta w$. Since

$$\begin{aligned} C(\breve{\mathbf{S}};\breve{p}) - C(\acute{\mathbf{S}};\acute{p}) &= \Delta g - \Delta w \\ &> 0 \end{aligned} \tag{4.34}$$

the expected cost of the least cost contract is higher to motivate $\hat{p} = \acute{p} - \epsilon$ than $\acute{p}$. The result holds for any ϵ and proves the proposition, so

$$\frac{\partial C(\mathbf{S}; P(\mathbf{S}))}{\partial p} < 0 \tag{4.35}$$

QED

Proposition 4.6 *The optimal cutoff point $\hat{p}$ is increasing in the manager's risk aversion, i.e. $\hat{p}_{\gamma_H} > \hat{p}_{\gamma_L}$ where $\gamma_H > \gamma_L$.*

Proof: The firm's objective function is maximized at $\hat{p}$ such that

$$\left.\frac{\partial V(p)}{\partial p}\right|_{p=\hat{p}} = \left.\frac{\partial C(\mathbf{S};p)}{\partial p}\right|_{p=\hat{p}} \tag{4.36}$$

$V(p)$ is a function of p only and independent of the risk aversion implementing the policy. To show that $V(p)$ is concave, note that

$$V(p) = pv_0 + \frac{1}{2}(1-p^2)v_g + \frac{1}{2}(1-p)^2 v_b \tag{4.37}$$

$$\frac{\partial V(p)}{\partial p} = v_0 - pv_g - (1-p)v_b \tag{4.38}$$

$$\frac{\partial^2 V(p)}{\partial p^2} = -(v_g - v_b) < 0 \tag{4.39}$$

Because $V(p)$ is concave, and we know that the optimal cutoff point $\hat{p}$ for any risk averse agent is less than the first best, p^*_{FB}, to prove the proposition it is sufficient to show that

$$\frac{\partial^2 C(\mathbf{S}; P(\mathbf{S}))}{\partial p \partial \gamma} < 0 \tag{4.40}$$

at every point $p < p^*_{FB}$.

Consider two agents, type L with risk aversion γ_L, and type H with risk aversion $\gamma_H = \gamma_L + \epsilon$. Consider the the pair of least cost contracts $\acute{\mathbf{S}}_L = (\acute{w}_L, \acute{g}_L), \acute{\mathbf{S}}_H = (\acute{w}_H, \acute{g}_H)$ that motivate investment policy $\acute{p}$ for both agents:

$$P_L(\acute{\mathbf{S}}_L) = P_H(\acute{\mathbf{S}}_H) = \acute{p} \tag{4.41}$$

Consider the cost of motivating either agent to reduce the investment cutoff point to $\breve{p} = \acute{p} - \delta$. Let $\breve{\mathbf{S}}_L = (\breve{w}_L, \breve{g}_L), \breve{\mathbf{S}}_H = (\breve{w}_H, \breve{g}_H)$ be the pair of least cost contracts that motivate the type L and type H agent, respectively, to pursue the new investment policy. By definition, these contracts satisfy

$$P_L(\breve{\mathbf{S}}_L) = P_H(\breve{\mathbf{S}}_H) = \breve{p} \tag{4.42}$$

Define Δg_L and Δg_H as the increase in the expected bonus payment under each of the two pairs contracts:

$$\Delta g_L = \frac{1}{2}(1 - \acute{p}^2)(\breve{g}_L - \acute{g}_L) + \frac{1}{2}[(1 - \breve{p}^2) - (1 - \acute{p}^2)]\breve{g}_L \tag{4.43}$$

$$\Delta g_H = \frac{1}{2}(1 - \acute{p}^2)(\breve{g}_H - \acute{g}_H) + \frac{1}{2}[(1 - \breve{p}^2) - (1 - \acute{p}^2)]\breve{g}_H \tag{4.44}$$

Let Δw_L and Δw_H be the difference in base pay under each of the two pairs of contracts:

$$\Delta w_L = \acute{w}_L - \breve{w}_L \tag{4.45}$$

$$\Delta w_H = \acute{w}_H - \breve{w}_H \tag{4.46}$$

Let ΔC_L and ΔC_H be the difference in cost for each of the two pairs of contracts:

$$\Delta C_L = \Delta g_L - \Delta w_L \tag{4.47}$$

$$\Delta C_H = \Delta g_H - \Delta w_H \tag{4.48}$$

From 4.13, $\frac{\partial g}{\partial \gamma} > 0$ and $\frac{\partial^2 g}{\partial \gamma \partial p} < 0$, so

$$\acute{g}_L < \acute{g}_H \tag{4.49}$$

$$\breve{g}_L < \breve{g}_H \tag{4.50}$$

$$(\breve{g}_L - g_L) < (\breve{g}_H - g_H) \tag{4.51}$$

$$\Delta g_L < \Delta g_H \tag{4.52}$$

For agent H, accepting $\breve{\mathrm{S}}_H$ and pursuing investment policy $\breve{p}$ is the same as accepting $\acute{\mathrm{S}}_H$, pursuing $\acute{p}$, and paying Δw_H for a lottery with expected payout Δg_H. Because $\acute{\mathrm{S}}_H, \breve{\mathrm{S}}_H$ are minimum cost contracts, both satisfy type H's reservation utility exactly:

$$\Psi_H(\acute{\mathrm{S}}_H; \acute{p}) = \Psi_H(\breve{\mathrm{S}}_H; \breve{p}) \tag{4.53}$$

Then it must be that

$$\Psi_H(\Delta g_H) = U_H(\Delta w_H) \tag{4.54}$$

From concavity of the utility function, for this equation to hold, $\Delta g_H > \Delta w_H$, and therefore

$$\begin{aligned} \Delta C_H &= \Delta g_H - \Delta w_H \\ &> 0 \end{aligned} \tag{4.55}$$

Define $\Delta \breve{w}_L$ as the amount agent L is willing to pay for a lottery paying the distribution Δg_H:

$$\Psi_L(\Delta g_H) = U_L(\Delta \breve{w}_L) \tag{4.56}$$

Because type H's utility function is a concave transformation of type L's, type L assigns a higher valuation to the uncertain payment, so

$$\Delta \breve{w}_L > \Delta w_H \tag{4.57}$$

Define $\Delta \check{C}_L$ as

$$\Delta \check{C}_L = \Delta g_H - \Delta \breve{w}_L \tag{4.58}$$

Because $\Delta C_H = \Delta g_H - \Delta w_H$

$$\Delta \check{C}_L < \Delta C_H \tag{4.59}$$

Now consider the reduction in base wage the type L agent takes moving from $\acute{\mathbf{S}}_L$ to $\check{\mathbf{S}}_L$:

$$\Psi_L(\Delta g_L) = U_L(\Delta w_L) \tag{4.60}$$

From equations 4.56 and 4.60, and by concavity of the utility function, it must be that

$$\Delta \breve{w}_L - \Delta w_L < \Delta g_H - \Delta g_L$$

because the certainty equivalent amount type L is willing to surrender in moving from lottery Δg_H (the left hand side of 4.56) to lottery Δg_L (the left hand side of 4.60) is less than the difference in expected payout of the two lotteries. Therefore

$$\Delta C_L < \Delta \check{C}_L < \Delta C_H$$

This holds $\forall\ p > p^*_{FB}, \epsilon, \delta$, therefore $\frac{\partial^2 C(\mathbf{S};P(\mathbf{S}))}{\partial p \partial \gamma} < 0$. This is sufficient to prove the proposition. Because this holds at any point $p > p^*_{FB}$, the slope of $\frac{\partial C(\mathbf{S};P(\mathbf{S}))}{\partial p}$ becomes steeper at any point p as γ increases. Because $V(p)$ is concave, the point at which $\left.\frac{\partial V(p)}{\partial p}\right|_{p=\hat{p}} = \left.\frac{\partial C(\mathbf{S};p)}{\partial p}\right|_{p=\hat{p}}$ is increasing in γ. *QED*

Proposition 4.7 *The equilibrium expected profit $\hat{\pi}(\hat{S};\hat{p}))$ is decreasing in γ.*

Proof: Consider two agents, L with risk aversion coefficient γ_L, H with risk aversion coefficient $\gamma_H = \gamma_L + \epsilon$. Let $\hat{p}_i$, $i = \{L, H\}$ be the equilibrium profit maximizing investment policy for agent i. From the previous proposition, $\hat{p}_L < \hat{p}_H$. Consider the profit from motivating either agent to pursue investment policy $\hat{p}_H$. Let $\hat{\mathbf{S}}_H$ be the contract that satisfies $\hat{p}_H = P_H(\hat{\mathbf{S}}_H)$, and define $\acute{\mathbf{S}}_L$ as the contract which satisfies $\hat{p}_H = P_L(\acute{\mathbf{S}}_L)$. Then the expected profit from each of the contractual relationships to motivate the investment policy is $\hat{p}_H$ given by

$$\pi(\hat{\mathbf{S}}_H;\hat{p}_H) = V(\hat{p}_H) - C(\hat{\mathbf{S}}_H;\hat{p}_H) \tag{4.61}$$

$$\pi(\acute{\mathbf{S}}_L;\hat{p}_H) = V(\hat{p}_H) - C(\acute{\mathbf{S}}_L;\hat{p}_H) \tag{4.62}$$

Because the expected cost of the wage contract motivating any investment policy p

is increasing in risk aversion,

$$C(\hat{\mathbf{S}}_H; \hat{p}_H) > C(\acute{\mathbf{S}}_L; \hat{p}_H) \tag{4.63}$$

Therefore

$$\pi(\acute{\mathbf{S}}_L; \hat{p}_H) > \pi(\hat{\mathbf{S}}_H; \hat{p}_H) \tag{4.64}$$

The profit maximizing contract for the type L agent is $\hat{\mathbf{S}}_L$ and satisfies $\hat{p}_L = P_L(\hat{\mathbf{S}}_L)$. By definition

$$\pi(\hat{\mathbf{S}}_L; \hat{p}_L) > \pi(\acute{\mathbf{S}}_L; \hat{p}_H) \tag{4.65}$$

Therefore

$$\pi(\hat{\mathbf{S}}_L; \hat{p}_L) > \pi(\hat{\mathbf{S}}_H; \hat{p}_H) \tag{4.66}$$

This holds for any ϵ, proving the proposition. *QED*

4.2.5 Discussion

The solution to the single agent case demonstrates that the agent's risk aversion is important from the perspective of the firm's shareholders. It also demonstrates that low risk aversion agents are desirable for two reasons. The first is the *wage cost effect.* The cost of motivating a given investment policy is increasing in the agent's risk aversion. This occurs because correct effort and project selection choices can only be motivated by risky pay. The cost of this is the deviation from optimal risk sharing. Risk averse agents look at a contract and demand that the expected utility it offers meet their level of reservation utility. Firm shareholders are risk neutral, and are only interested in the expected cost of the wage package. The more risk averse the manager, the more costly it becomes for the firm to offer a risky pay package meeting a given level of expected utility.

The second reason that low risk aversion agents are more desirable is the *project selection effect*. As the agent's risk aversion increases, the best investment policy that the agent can profitably be persuaded to follow is increasingly distorted from the first best case. This distortion takes the form of underinvestment, meaning that risky projects are rejected that the shareholders would ideally prefer a perfectly aligned manager to pursue. This occurs because as the agent's risk aversion increases, his risk preferences are increasingly different from those of risk neutral shareholders.

4.3 Model: Competitive Labour Market, Two Agent Types

Consider the same problem, but in the context of a labour market where the shareholders of n identical firms must compete with each other for the services of managers. These managers are identical in every respect, with the exception of their risk aversion. There are m agents of type L with low risk aversion γ_L, while the remaining agents have risk aversion γ_H, $\gamma_H > \gamma_L$. m is less than n, so all m type L agents are employed. The remaining $n - m$ jobs are filled by agents of type H.

An agent's type is private information, and cannot be communicated to firm shareholders. As such, firms offer incentive compatible contracts $[\mathbf{S}_L, \mathbf{S}_H]$ that leads agents to truthfully identify their type through the contract they choose.

However, firms do not operate in isolation. They operate in a competitive labour market, where all firms will bid for the services of desirable low risk aversion (type L) managers. The equilibrium concept is a Nash equilibrium in contract offers.

A Nash equilibrium in this market takes the following form: m firms offer contract $\mathbf{S}_L$ and hire a type L agent, $n-m$ firms offer contract $\mathbf{S}_H$ and hire a type H agent, and no firm has an incentive to deviate by offering some other contract $\mathbf{S}'$. One condition of such an equilibrium is that

$$\pi(\mathbf{S}_L; p_L) = \pi(\mathbf{S}_H; p_H)$$

4.3.1 Full information case

To illustrate the importance screening plays in the equilibrium, first consider the equilibrium in a case where agents' type is observable. Each firm can offer a contract designed for a given type of agent. Both agents demand a certainty equivalent wage of $\bar{w}$ to participate in the game. The equilibrium investment cutoff point, $\hat{p}_i$, is the same for each agent as it would be in the single agent case. The investment cutoff is the one which satisfies

$$\left.\frac{\partial V(p)}{\partial p}\right|_{p=\hat{p}_i,i} = \left.\frac{\partial C(\mathbf{S};p)}{\partial p}\right|_{p=\hat{p}_i,i} \tag{4.67}$$

Contracts are of the form $\mathbf{S}_i = (w_i, g_i)$, and are the minimum cost contract that motivate the investment cutoff p_i. The propositions proved in the previous section show that in equilibrium, type L agents pursue a more aggressive investment policy, $\hat{p}_L < \hat{p}_H$. When each agent is held at his reservation utility, it is more profitable to contract with type L agents. If there were more type L agents than firms in the market $(m > n)$, then there would be no role for type H agents. Since $\pi(\hat{\mathbf{S}}_L; \hat{p}_L) > \pi(\hat{\mathbf{S}}_H; \hat{p}_H)$ when the incentive compatibility constraint binds for both types, hiring a type L agent is more profitable, and all firms would do so.

However, when there is a shortage of type L agents relative to the total number of firms needing agents ($m < n$, as is assumed to be the case), type H agents have a role to play in the labour market. If firms were able to contract either type of agent at their reservation level of utility, then it would be profitable to choose type L agents. Any firm contracting with type H would have an incentive to deviate from such an equilibrium offering a contract $\breve{\mathbf{S}}_L = (\hat{w}_L + \delta, \hat{g}_L)$. This deviation increases type L's expected utility by increasing the fixed payment by δ, while leaving the bonus $\hat{g}_L$, and therefore the investment policy $\hat{p}_L$, unchanged.

Because in equilibrium firms must not have incentive to deviate, the pair of contracts $\hat{\mathbf{S}}_L, \hat{\mathbf{S}}_H$ that would each be optimal in a single agent framework cannot constitute an equilibrium. Firms contracting with type L agents offer an increase t to the fixed wage, defined as:

$$t = \pi(\hat{\mathbf{S}}_L; \hat{p}_L) - \pi_H(\hat{\mathbf{S}}_H; \hat{p}_H) \tag{4.68}$$

where $\hat{\mathbf{S}}_L$ is the least cost contract motivating type L to follow $\hat{p}_L$ in the single agent case, and $\hat{\mathbf{S}}_H$ motivates type H to follow $\hat{p}_H$.

The equilibrium contract menu is then

$$\begin{aligned} \acute{\mathbf{S}}_L &= (\hat{w}_L + t, \hat{g}_L) \\ \hat{\mathbf{S}}_H &= (\hat{w}_H, \hat{g}_H) \end{aligned} \tag{4.69}$$

These contracts motivate $\hat{p}_L, \hat{p}_H$ respectively, and

$$\pi(\acute{\mathbf{S}}_L; \hat{p}_L) = \pi(\mathbf{S}_H; \hat{p}_H) \tag{4.70}$$

Type H agents are kept strictly at their reservation level of utility.[11] Since type L agents create more asset value and are in scarce supply, they earn rents in equilibrium:

$$\Psi(\acute{\mathbf{S}}_L; \hat{p}_L) > \Psi(\hat{\mathbf{S}}_L; \hat{p}_L) = U_L(\bar{w}) \tag{4.71}$$

In this full information case, there is no need to worry about self-selection constraints on the pair of contracts. The rents paid to type L come in the form of increased base pay, t. The level of investment under the contract for type L, $\hat{p}_L$, is the same as it would be if only type L agents were present in the labour market, and the same holds true for investment under the contract for H.

[11] In a competitive market, this reservation utility $\bar{U} = U_H(\bar{w})$ is such that there is zero expected profit net of wages.

It is instructive to consider the nature of rents paid to type L in this case. These rents are not informational rents, since this is a full information case. Rather, these are Ricardian rents accruing to type L because of their value (from their ability to generate a higher expected firm value) and their scarcity in the market.

4.3.2 Private information case

The game changes when the agent's type is private information. In equilibrium, firms offer screening contracts that elicit truthful revelation of type. As such, the contract for type i must not only satisfy the constraints of the single agent case, but also self-selection constraints. These prevent one type from mimicking the other and choosing the contract designed for the other agent.

The expected profit for a firm contracting with an agent of type i using contract $\mathbf{S}_i = (w_i, g_i)$ to motivate $p_i = P_i(\mathbf{S}_i)$ is

$$\pi(\mathbf{S}_i; p_i)) = V(p_i) - C(\mathbf{S}_i; p_i) \tag{4.72}$$

where

$$V(p_i) = p_i v_0 + (1 - p_i^2)v_g + (1 - p_i)^2 v_b \tag{4.73}$$

$$C(\mathbf{S}_i; p_i) = w_i + (1 - p_i^2)g_i \tag{4.74}$$

Define agent i's expected utility from pursuing investment policy p on contract $\mathbf{S}$ as

$$\Psi_i(\mathbf{S};p) = pU_i(w) + \frac{1}{2}(1 - p^2)U_i(w + g - c) + \frac{1}{2}(1 - p)^2 U_i(w - c)$$

Then in equilibrium contracts $\hat{\mathbf{S}}_i = \{\hat{\mathbf{S}}_L, \hat{\mathbf{S}}_H\}$ must satisfy

$$\Psi_i(\hat{\mathbf{S}}_i;\hat{p}_i) \geq U_i(\bar{w}) \tag{4.75}$$

$$\hat{p}_i = P_i(\hat{\mathbf{S}}_i) = \frac{U_i(0) - U_i(-c)}{U_i(\hat{g}_i - c) - U_i(-c)} \tag{4.76}$$

$$\Psi_i(\hat{\mathbf{S}}_i;\hat{p}_i) \geq \Psi_i(\hat{\mathbf{S}}_j;\check{p}_i = P_i(\hat{\mathbf{S}}_j)) \tag{4.77}$$

with $[i, j] = [L, H]$, $i \neq j$.

4.75 is the incentive compatibility constraint, and assumes that each agent has the same reservation certainty equivalent wage, $\bar{w}$. Equation 4.76 defines a type i agent's individually rational investment cutoff decision for a given contract.

Equation 4.77 is a self-selection constraint which requires that agent i prefer contract $\hat{\mathbf{S}}_i$, that intended for his own type, rather than the contract intended for the other agent. $\Psi_i(\mathbf{S}_j;\check{p}_i = P_i(\mathbf{S}_j))$ represents the highest expected utility an agent of type i could obtain by taking the contract intended for the other agent type.

Each firm maximizes its expected value, subject to the contracts offered by other firms in the competitive market for managerial labour. In equilibrium, firms must have no incentive to deviate, and must therefore be indifferent as to whether they contract with type L or type H agents. Therefore

$$\pi(\hat{\mathbf{S}}_L; \hat{p}_L) = \pi(\hat{\mathbf{S}}_H; \hat{p}_H) \tag{4.78}$$

in equilibrium. This equation holds that expected asset value less expected wage costs must be the same for either agent type in equilibrium.

Before moving to the formal characterization of the equilibrium, it is helpful to discuss the possible outcomes heuristically. First, type H agents receive the same contract as they would in the single agent case. Because type H agents add less firm value than type L agents, firms do not have an incentive to bid for their services beyond the basic level of utility they would receive in the single agent case. Since firms have no incentive to change the contract they offer to type H agents, $\hat{\mathrm{S}}_H$ is the same contract as would be offered without the introduction of a second (more desirable) type of agent.

Type L agents will receive rents, as demonstrated in the full information case. What may change, depending on the parameters, is the way in which those rents are paid. Relative to the single agent type case, type L agents receive rents based on their ability to add more firm value than type H. Because in equilibrium firms must be indifferent between hiring either type of agent, and type L agents create higher valued firms gross of expected wage costs, expected wage costs, and ultimately expected utility, are higher for type L.

However, the manner in which type L agents receive their rents depends on the relative value type H agents place on their own contract, $\hat{\mathrm{S}}_H$, and the contract intended for type L, $\hat{\mathrm{S}}_L$. As long as type H agents prefer their own contract, firms can pay rents to type L agents in the most efficient way possible: increase base pay on $\hat{\mathrm{S}}_L$. If there is some level of increased base pay, without changing type L bonus pay, at which firms are indifferent between agent types, then type L agents receive rents in the form of increased base pay only, and there is no change in investment policy.

However, if base pay to type L increases to the point that type H agents would start to prefer the type L contract if base pay increases further, and firms still prefer type L agents, then rents paid to type L agents must take a second form: increased *bonus* pay, in addition to the increased base pay. This solution has the advantage that the more risk averse type H agents find the bonus pay less attractive than do type L agents, thereby performing the screening function required of the contract menu. However, increased bonus pay changes type L investment policy from the single agent case, and is an inefficient means of paying rents relative to the full infromation solution. Because the rents come in the form of risky pay, risk averse agents value the bonus payments less than they do certain payments.

Properties of the equilibrium

Lemma 4.2 *The equilibrium wage contract designed for the type H agent, $\hat{S}_H$, is the same contract that would be offered in a labour market populated by type H agents only.*

Proof: As in the single agent case, the incentive compatibility constraint (4.75) must bind for the type H agent. Because it is more costly to motivate type H agents to pursue any investment policy p, there is no reason for firms to work to attract a type H agent over a type L agent. As such, firms contracting with type H agents design the contract which maximizes expected profit. This is the investment cutoff $\hat{p}_H$ that satisfies

$$\left.\frac{\partial V(p)}{\partial p}\right|_{p=\hat{p}_H} = \left.\frac{\partial C(\mathbf{S}_H;p)}{\partial p}\right|_{p=\hat{p}_H} \tag{4.79}$$

These are exactly the same equilibrium conditions as in the single agent case, so $\hat{\mathbf{S}}_H$, is unaffected by the presence of the type L agent in a labour market setting. *QED*

Proposition 4.8 *Type L agents earn rents in equilibrium, and the incentive compatibility constraint (equation 4.75) does not bind for type L agents.*

Proof: Consider a contradiction. Let $\acute{\mathbf{S}}_L$ be the equilibrium wage contract, intended for type L agents, offered by m firms. Let $\hat{\mathbf{S}}_H$ be the contract offered by $n-m$ firms, intended for type H agents. If the type L agent's incentive compatibility constraint binds, then

$$\Psi_L(\acute{\mathbf{S}}_L;\acute{p}_L) = U_L(\bar{w}) \tag{4.80}$$

From the previous proposition, we know that type H's incentive compatibility constraint binds:

$$\Psi_H(\hat{\mathbf{S}}_H;\hat{p}_H) = U_H(\bar{w}) \tag{4.81}$$

From the previous section, the type L agent can generate a higher certainty equivalent utility than type H on any contract. Therefore

$$CE_L(\Psi_L(\hat{\mathbf{S}}_H;\check{p}_L = P_L(\mathbf{S}_H))) > CE_H(\Psi_H(\hat{\mathbf{S}}_H;\hat{p}_H)) = \bar{w} \tag{4.82}$$

$$> CE_L(\Psi_L(\acute{\mathbf{S}}_L;\acute{p}_L)) \tag{4.83}$$

$$\Psi_L(\hat{\mathbf{S}}_H;\check{p}_L) > \Psi_L(\acute{\mathbf{S}}_L;\acute{p}_L) \tag{4.84}$$

Equations 4.82 - 4.84 show that $\acute{\mathbf{S}}_L$ violates the self-selection constraint 4.77. Therefore, type L's incentive compatibility constraint cannot bind in equilibrium, and type L agents earn rents. *QED*

Proposition 4.9 *If for an equilibrium pair of contracts $\hat{S}_L$,$\hat{S}_H$ the self-selection constraint 4.77 does not bind for type H agents, then $\hat{S}_L$ motivates the same investment policy $\hat{p}_L$ as would be optimal in a labour market populated by type L agents only.*

Proof: Let $\hat{\mathbf{S}}_H$ be the equilibrium contract offered to type H agents. Let $\hat{\mathbf{S}}_L = (\hat{w}_L, \hat{g}_L)$ be the contract that motivates the same investment policy, $\hat{p}_L$, as would be optimal in the single-agent case populated only by type L agents, and satisfies the self-selection constraints so that

$$\Psi_L(\hat{\mathbf{S}}_L;\hat{p}_L) \;>\; \Psi_L(\hat{\mathbf{S}}_H;\check{p}_L = P_L(\hat{\mathbf{S}}_H)) \tag{4.85}$$

$$\Psi_H(\hat{\mathbf{S}}_H;\hat{p}_H) \;>\; \Psi_H(\hat{\mathbf{S}}_L;\check{p}_H = P_H(\hat{\mathbf{S}}_L)) \tag{4.86}$$

as well as the equilibrium profit condition

$$\pi(\hat{\mathbf{S}}_L;\hat{p}_L) = \pi(\hat{\mathbf{S}}_H;\hat{p}_H) \tag{4.87}$$

Since $\hat{\mathbf{S}}_L$ motivates the single agent case investment policy,

$$\left.\frac{\partial V(p)}{\partial p}\right|_{p=\hat{p}_L,L} = \left.\frac{\partial C(\mathbf{S}_H;p)}{\partial p}\right|_{p=\hat{p}_L,L} \tag{4.88}$$

Let $\acute{\mathbf{S}}_L$ be a contract that motivates a different investment policy, $\acute{p}_L$, and keeps the type L agent's utility unchanged:

$$\Psi_L(\hat{\mathbf{S}}_L;\hat{p}_L) = \Psi_L(\acute{\mathbf{S}}_L;\acute{p}_L) \tag{4.89}$$

where

$$\acute{\mathbf{S}}_L = (\hat{w}_L - f(\delta), \hat{g}_L + \delta) \tag{4.90}$$

Because $P_L(\hat{\mathbf{S}}_L) \neq P_L(\acute{\mathbf{S}}_L)$, any δ implies lower expected profit for the firm, since

$$\frac{\partial V(p)}{\partial p}_{p=\acute{p}_L,L} \neq \frac{\partial C(\mathbf{S}_L;p)}{\partial p}_{p=\acute{p}_L,L} \tag{4.91}$$

Because this is true for all δ, there exists no contract which is a Pareto improvement over $\hat{\mathbf{S}}_L$, which proves the proposition. *QED*

Proposition 4.10 *Let $\acute{p}_L$ be the level of investment motivated by optimal contract for a labour market populated by type L agents only, and let $\acute{g}_L = G_L(\acute{p}_L)$ be the bonus payment that motivates the type L agent to pursue policy $\acute{p}_L$. Let $\hat{S}_L$,$\hat{S}_H$ be the equilibrium contracts. If there exists a contract $\acute{S}_L = (\acute{w}_L, \acute{g}_L)$ for which (i) the type H agent's self-selection constraint binds, and (ii) for which $\pi(\acute{S}_L;\acute{p}_L) > \pi(\hat{S}_H;\hat{p}_H)$ then the equilibrium contract for type L motivates a more aggressive investment policy than in the single agent case.*

Proof: $\acute{\mathbf{S}}_L, \hat{\mathbf{S}}_H$ cannot be an equilibrium as $\pi(\acute{\mathbf{S}}_L;\acute{p}_L) > \pi(\hat{\mathbf{S}}_H;\hat{p}_H)$. The equilibrium type L contract, $\hat{\mathbf{S}}_L$, must satisfy $\pi(\hat{\mathbf{S}}_L;\hat{p}_L) = \pi(\hat{\mathbf{S}}_H;\hat{p}_H)$. Therefore $\pi(\hat{\mathbf{S}}_L;\hat{p}_L) < \pi(\acute{\mathbf{S}}_L;\acute{p}_L)$.

Because

$$\Psi_H(\hat{\mathbf{S}}_H;\hat{p}_H) = \Psi_H((\acute{\mathbf{S}}_L;\acute{p}_H = P_H(\acute{\mathbf{S}}_L)) \tag{4.92}$$

then for any contract $\breve{\mathbf{S}}_L = (\acute{w}_L + \epsilon, \acute{g}_L)$ it will be the case that

$$\Psi_H(\hat{\mathbf{S}}_H; \hat{p}_H) < \Psi_H(\breve{\mathbf{S}}_L; \acute{p}_H) \tag{4.93}$$

as the base wage is higher for $\breve{\mathbf{S}}_L$ but the bonus payment is the same for both contracts. This makes the agent strictly better off, and violates the self-selection constraint 4.77 for agent H. Then the equilibrium contract $\hat{\mathbf{S}}_L$ must be such that $\hat{w}_L < \acute{w}_L$.

The equilibrium contract menu $\hat{\mathbf{S}}_L$,$\hat{\mathbf{S}}_H$ must therefore satisify

$$\Psi_L(\hat{\mathbf{S}}_L; \hat{p}_L) \geq \Psi_L(\acute{\mathbf{S}}_L; \acute{p}_L) \tag{4.94}$$

$$\Psi_H(\hat{\mathbf{S}}_H; \hat{p}_H) \leq \Psi_H(\hat{\mathbf{S}}_L; \bar{p}_H = P_H(\hat{\mathbf{S}}_L)) \tag{4.95}$$

$$\pi(\hat{\mathbf{S}}_L; \hat{p}_L) = \pi(\hat{\mathbf{S}}_H; \hat{p}_H) \tag{4.96}$$

For any contract $\grave{\mathbf{S}}_L = (\acute{w}_L - \epsilon, \acute{g}_L)$, $\epsilon > 0$,

$$\Psi_L(\grave{\mathbf{S}}_L; \acute{p}_L) < \Psi_L(\hat{\mathbf{S}}_L; \hat{p}_L) \tag{4.97}$$

Because $\hat{w}_L < \acute{w}_L$, by equation 4.94:

$$\hat{g}_L > \acute{g}_L \tag{4.98}$$

Because $\hat{g}_L > \acute{g}_L$, and $\frac{\partial p}{\partial g} < 0$, $\hat{p}_L > \acute{p}_L$.

Finally, for there to exist an equilibrium contract $\hat{\mathbf{S}}_L$, it must satisfy

$$\Psi_H(\hat{\mathbf{S}}_L; \bar{p}_H = P_H(\hat{\mathbf{S}}_L)) = \Psi_H(\acute{\mathbf{S}}_L; \acute{p}_H = P_H(\acute{\mathbf{S}}_L)) \tag{4.99}$$

$$\Psi_L(\hat{\mathbf{S}}_L; \hat{p}_L) > \Psi_L(\acute{\mathbf{S}}_L; \acute{p}_L) \tag{4.100}$$

Let $\Delta w_L = \acute{w}_L - \hat{w}_L$, and $\Delta g_L = \hat{g}_L - \acute{g}_L$. $\hat{\mathbf{S}}_L$ is equivalent to $\acute{\mathbf{S}}_L$ plus paying Δw_L for a lottery paying Δg_L in the event the good state is realized. From 4.99, it must be that

$$CE_H(\Psi_H(\Delta g_L) - \Delta w_L = 0 \tag{4.101}$$

Because type L can generate higher certainty equivalent utility from any lottery

$$CE_L(\Psi_L(\Delta g_L) - \Delta w_L > 0 \tag{4.102}$$

Therefore a contract satisfying equations 4.94 to 4.96, 4.99 and 4.100 can be found. This proves the proposition. *QED*

Corollary 4.2 *When equation 4.77 binds for type H agents, the equilibrium contract menu may motivate type L agents to pursue a policy of overinvestment relative to the first best case.*

This is a natural extension of the previous proposition. A parametric example is sufficient to demonstrate the existence of such equilibria.

Consider an example where $\{v_B = 0, v_0 = 100, v_G = 400, c = 2, \gamma_L = 0.1, \gamma_H = 0.5\}$ and both agents demand a utility equal to that provided by a certainty equivalent wage of $\bar{w} = 5$. The optimal contract to be offered to the type H agent is $\hat{\mathbf{S}}_H = (\hat{w}_H, \hat{g}_H) = (4.606, 10.179)$. This maximizes expected profit $\pi(\hat{\mathbf{S}}_H; \hat{p}_H = P_H(\hat{\mathbf{S}}_H) = 0.636) = 175.058$.

The equilibrium contract $\hat{\mathbf{S}}_L$ that satisfies both the type H self-selection constraint and the equilibrium labour market requirement that $\pi(\hat{\mathbf{S}}_L;\hat{p}_L) = \pi(\hat{\mathbf{S}}_H;\hat{p}_H)$ is $\hat{\mathbf{S}}_L(\hat{w}_L,\hat{g}_L) = (4.593, 65.995)$. Type H's self-selection constraint is satisfied; $\Psi_H(\hat{\mathbf{S}}_H;\hat{p}_H) = \Psi_H(\hat{\mathbf{S}}_L;\bar{p}_H = P_H(\hat{\mathbf{S}}_L) = 0.632) = -0.082$. The expected profit when the type L agent takes the contract is $\pi(\hat{\mathbf{S}}_L;\hat{p}_L = P_L(\hat{\mathbf{S}}_L) = 0.182) = 175.058$, the same as the expected profit when contracting with a type H agent on $\hat{\mathbf{S}}_H$.

In a full-information environment with observable effort, the first-best level of investment is given by $p^*_{FB} = \frac{v_0 - v_b + c}{v_g - v_b} = 0.245$. Comparing this cutoff point to the equilibrium cutoff point of $\hat{p}_L = 0.182$, it is clear that in this equilibrium type L agents are provided incentives to overinvest relative to the first-best case.

Proposition 4.10 *In equilibrium, firms contracting with type L agents have higher ex ante variance of expected firm value than do firms contracting with type H agents.*

Proof: The ex ante distribution of firm value, after the agent is hired but before the updated signal $\tilde{r}$ is received, is trinomial. When the investment cutoff is p expected firm value is given

$$E(V) = pv_0 + (1-p^2)v_g + (1-p)^2 v_b \tag{4.103}$$

The variance of firm value is given

$$Var(V) = E(V^2) - [E(V)]^2 \tag{4.104}$$

whose first derivative with respect to p is given

$$[-(1-p)v_b^2 - pv_g^2 + v_0^2] - [(-1+p)v_b - pv_g + v_0][(1-p)^2 v_b + v_g - p^2 v_g + 2pv_0] \quad (4.105)$$

This expression is negative for all $p \in (0,1)$ where $v_b < v_0 < v_g$. Because $\hat{p}_L < \hat{p}_H$, the ex ante variance of firm value is greater for firms hiring type L agents. *QED*

Discussion

There are two key changes in the case where multiple firms compete for the services of two different types of agents. The first is that firms introduce a menu of contracts, and these contracts perform a screening function to distinguish between agent types. The second is that because firms have an incentive to bid against each other for the scarce services of type L managers, type L managers capture rents in equilibrium.

The nature of rents in this model differs from that of most screening models. Typically, the "good" type in a given model earns rents due to the self-selection constraint. The desirable agent receives rents in order to elicit truthful revelation of type. A portion of the rents the desirable type L agent receives in the model comes from this source. Firms cannot offer the type L agent the same contract they would in a single-agent setting, because the type L contract can do better by accepting the contract intended for type H.

However, this is not the primary source of rents to the agent in the model. Rents accrue to type L agents because their lower risk aversion leads them to make less distorted investment choices relative to a first-best (or risk neutral agents) case. In a labour market where firms have the ability to bid up the prices of agents, type L agents capture the extra value that they create. As long as the type H agents' self-selection constraint does not bind, increased pay to type L agents takes the form of increased base pay. This is the most efficient type of payment from a risk sharing perspective.

The interesting cases are those where the high type's self-selection constraint is binding. The nature of the equilibrium is much different than the result in a single-agent case. Here, firms contracting with type L agents must design a contract which performs a screening function, in that it has to satisfy type H agents' self-selection constraint. At the same time, in equilibrium the contract provides rents to type L to the degree that firms are indifferent between contracting with either type of agent. Put another way, type L agents capture all of the extra value they create.

The solution which satisfies both the screening requirement and the labour market requirement is one where rents to type L take the form of far more bonus compensation than they receive in a single agent type setting. The use of bonus pay takes advantage of the different agent types' feelings about risk. Type L agents value the bonus compensation more highly and earn rents by taking the contract. Type H agents assign a much lower value to this riskier contract, and as a result their self-selection constraint is satisfied.

The interesting result that emerges from this labour market is that the presence of type H agents can cause firms to increase incentive pay to type L agents beyond what they would receive in a single-agent market. The contracts for type L agents can exhibit reduced project selection distortion relative to the single agent case. Increasing risky pay leads to a more aggressive investment policy, reducing the degree of underinvestment that arises in the single-agent case. However, this comes at the cost of increased wages and poorer risk sharing than in the single-agent case. Clearly, the wage cost and risk sharing effects must outweigh the value of the improved project selection, otherwise the level of investment motivated in the dual-agent case would also have been optimal for the single-agent case.

The most interesting equilibrium is in cases where type L agents actually overinvest relative to the first best case. This provides a stark example of just how important the contracts' screening role can be in the two agent case. Risky pay, expensive due to the deviation from optimal risk sharing that is required to make it work, is used to the point that type L agents choose to overinvest. The need for screening in these cases causes firms to offer a risk averse agent a contract of higher power than they would offer even a risk neutral agent, with whom risk sharing concerns would be irrelevant.

The intuition for this relates to the reason that type L agents earn rents in the model. Rather than being paid rents to ensure that they don't mimic type H, the rents paid to type L have more to do with a traditional labour supply and labour demand curves. There are relatively few type L agents, they add more value, and in equilibrium they are paid more to reflect this. This causes equilibria where the type H agent's self-selection constraint binds instead. This is the reason that rents paid to type L take a form (bonus pay) that is inefficient from the perspective of straight risk sharing-project selection tradeoff concerns. The bonus is not valued by type L agents anywhere near its true value. In cases where overinvestment results, some of the "rents" are paid in the form of asset value destroyed by overinvestment. This is clearly not desirable, but it is a consequence of firms' efforts to avoid value being lost because type H agents choose to underinvest in equilibrium.

Finally, it is informative to compare the interaction between firm risk and risk sharing in this model with that of many traditional principal-agent models. Generally, in models where firms provide agents with risky pay in order to motivate effort, the amount of risky pay they offer is decreasing in the firm's (generally exogenously specified) firm risk. The expected cost of providing risk averse agents with incentive pay to provide them a given level of expected utility is increasing in firm risk. As the firm's returns become riskier, the tradeoff between the effort incentives from risky pay and the cost of the risky pay causes the firm to choose lower powered contracts. This contrasts with the empirical literature, where Prendergast finds that the link between firm risk and the level of risky pay provided is quite mixed.

In this model, firm risk is not an exogenous parameter around which firms make contracting choices. Rather, *firm risk is an endogenous outcome of the contracting choices a given firm makes.* In this model, agents influence firm risk because they have the discretion to make project selection decisions. Firms that contract with type L agents provide greater incentive pay than do firms that contract with type H agents. The higher power of type L contracts, combined with type L agents' lower risk aversion and project selection ability, leads firms managed by L-type agents to have higher ex ante variance in firm value than firms managed by type H agents. This is the exact opposite predicted by a model where managers do not influence project selection and firm risk is exogenous.

4.4 Conclusion

The paper develops a model where incentive contracts are designed to elicit effort, motivate properly aligned project selection and investment decisions, and screen potential candidates. I demonstrate the importance of screening in a setting where agents of differing risk aversion populate the labour market. Firms have an incentive to bid for low risk aversion agents, who can be encouraged to pursue a less distorted investment policy. The resulting labour market equilibrium leads to outcomes where contracts have much greater power than they do in versions of the same model where screening is not a consideration.

This result sheds new light on why very high power contracts are often observed empirically. Such contracts are very hard to justify using a traditional model which trades off costly risk sharing against the need to motivate effort. The introduction of a screening component to the model provides a reasonable justification.

This paper also incorporates the realistic feature that high-level managers do not only exert productive effort on a given project. They also crucially make investment and project selection decisions that affect the distribution of firm returns. In such a setting, managerial risk aversion becomes an extremely important consideration in contract design, since managers with different degrees of risk aversion make different investment decisions when faced with the same contract.

The results provide new understanding of the roles played by incentive compensation, as well as the optimal degree of incentive power provided by such contracts.

4.5 Bibliography

[1] Carpenter, Jennifer N., "Does Option Compensation Increase Managerial Risk Appetite?", *Journal of Finance*, October 2000, V.55 No.5, 2311-2331.

[2] Core, John E., Wayne Guay and David F. Larcker, "Executive Equity Compensation and Incentives: A Survey", 2002, working paper.

[3] Core, John E. and Jun Qian, "Project Selection, Production, Uncertainty, and Incentives", 2002, working paper.

[4] Grossman, Sanford J. and Oliver D. Hart, "An Analysis of the Principal-Agent Problem", *Econometrica,* January 1983, V.51 No.1, 7-46.

[5] Hall, Brian and Kevin J. Murphy, "Stock Options for Undiversified Executives", 2001, USC working paper 01-16.

[6] Harris, Milton and Artur Raviv, "Some Results on Incentive Contracts with Applications to Education and Employment, Health Insurance, and Law Enforcement", *American Economic Review,* March 1978, V.68 No. 1, 20-30.

[7] Haubrich, Joseph G., "Risk Aversion, Performance Pay, and the Principal-Agent Problem", *Journal of Political Economy,* April 1994, V.102 No. 2, 258-276.

[8] Holmstrom, Bengt, "Moral Hazard and Observability", Bell Journal of Economics, 1979, V.10 No. 1, 74-91.

[9] Holmstrom, Bengt, "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies*, 1982, V.66 No. 1, 169-182.

[10] Holmstrom, Bengt and Paul Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica,* March 1987, V.55 No. 2, 303-328.

[11] Jensen, Michael C. and William Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, October 1976, V.3 No. 4, 305-360.

[12] Jensen, Michael C. and Kevin J. Murphy, "Performance Pay and Top-Management Incentives", *Journal of Political Economy*, April 1990, V.98, 225-264.

[13] Jullien, Bruno, Bernard Salanie and Francois Salanie, "Screening Risk-Averse Agents Under Moral Hazard", 2000, working paper.

[14] Lambert, Richard. A., "Executive Effort and Selection of Risky Projects", *RAND Journal of Economics*, Spring 1986, V.17 No. 1, 77-88.

[15] Lewellen, Katharina, "Financing Decisions When Managers Are Risk Averse", 2003, MIT working paper 4438-03.

[16] Murphy, Kevin J., "Executive Compensation", 1999, working paper.

[17] Nohel, Tom and Steven Todd, "Stock Options and Managerial Incentives to Invest", 2001, working paper.

[18] Prendergast, Canice, "The Tenuous Trade-off between Risk and Incentives", *Journal of Political Economy*, 2002, V.110 No. 5, 1071-1102.

[19] Ross, Stephen A., "The Economic Theory of Agency: The Principal's Problem", *American Economic Review*, May 1973, V.63 No. 2, 134-139.

[20] Ross, Stephen A., "Compensation, Incentives, and the Duality of Risk Aversion and Riskiness", *Journal of Finance*, Feb. 2004, V.59 No. 1, 207-225.

[21] Rothschild, Michael and Joseph E. Stiglitz, "Increasing Risk: I. A Definition", *Journal of Economic Theory*, 1970, V.2 No. 3, 225-243.

[22] Serfes, Konstantinos, "Endogenous Matching in a Market with Heterogeneous Principals and Agents", 2005, working paper.

[23] Wright, Donald J., "The Risk and Incentives Trade-off in the Presence of Heterogeneous Managers", *Journal of Economics*, 2004, V.83 No. 2, 209-223.

Chapter 5

Conclusion

In the first essay I seek to explain how tort liability affects a firm's optimal capital structure. While other papers have made the point that limited liability will affect economic agents' incentives with respect to tort risk, very few have sought to endogenize the firm's decision about in which states it will be solvent. A key characteristic of tort risk is that its impact on cash flows available to security holders depends on the structure of security holders' claims. Put another way, capital structure matters greatly when determining the potential expense payable to tort claimants. Recognizing this, firms with exposure to tort liability will have an incentive to adjust capital structure to respond optimally. The lower creditor priority of tort claimants implies two effects when debt and tort risk interact. The first is that tort liability brings about an increased probability of bankruptcy. Where this effect predominates, the firm will choose to move away from debt. The second effect is that debt provides an asset shielding advantage, preserving cash flow rights for the firm's debt holders at the expense of tort claimants. Where this effect is dominant, increased tort risk will cause the firm to choose more debt.

I specify two simple models to examine the interaction of these effects, one where firm returns are distributed continuously over an interval, and another where firm returns are distributed binomially. The different results from these two illustrations demonstrate the importance of assumptions regarding firm cash flows. Depending on the nature of the firm's returns, and the values of the various input parameters, either the bankruptcy effect or asset shielding effect can dominate.

Tort liability is a major source of risk for firms today. I have shown why it is unique, and why firms must consider its unique properties when determining the optimal capital structure. Empirical work studying how firms do adjust their capital structure to address changes in tort risk is a potentially fruitful avenue for future research.

In the second essay we explore the capital structure of insurers. The focus is on the impact of aggregate uncertainty, or dependence among risks, since this is the source of an insurer's incentive to issue equity. Insurance firms respond to the shocks of increased risks by taking all or some of the following actions: placing limits on the number or coverage of contracts that they offer; raising premium for the policies that they issue; and raising more equity. We analyze the equilibrium mixture of these responses in a competitive insurance market, and find that the impact of increasing uncertainty on the equity decision depends on the nature of aggregate uncertainty. Where this uncertainty is in the size of losses, equity increases with uncertainty; where the risk dependence is in the events of losses, equity first increases then decreases with uncertainty, providing that individuals are not too risk averse. The latter result follows from a tradeoff between two effects, which we label the input effect of uncertainty, and the output effect. In both cases, however, the ratio of equity to insurance revenue increases. We extend the model to look at the effect of a cost difference between internal equity (less costly) and external equity (more costly). This extension leads to the hypothesis that firms with greater internal equity will tend to use less leverage.

We test both hypotheses directly on a sample of 852 U.S. property and casualty stock insurers over a sample period from 1999-2004. We find support for both of our hypotheses. Firms that have higher variance in their loss ratio, our proxy for uncertainty, use significantly less leverage, supporting our theory that uncertainty and leverage are negatively correlated. Firms that have been recently profitable, implying greater internal capital, use significantly less leverage. This supports the theory that there is a cost advantage to internal over external equity, which is at the core of recent theories of insurance market dynamics.

In the third essay, I develop a model where incentive contracts are designed to elicit effort, motivate properly aligned project selection and investment decisions, and screen potential candidates. I demonstrate the importance of screening in a setting where agents of differing risk aversion populate the labour market. Firms have an incentive to bid for low risk aversion agents, who can be encouraged to pursue a less distorted investment policy. The resulting labour market equilibrium leads to outcomes where contracts have much greater power than they do in versions of the same model where screening is not a consideration.

This result sheds new light on why very high power contracts are often observed empirically. Such contracts are very hard to justify using a traditional model which trades off costly risk sharing against the need to motivate effort. The introduction of a screening component to the model provides a reasonable justification.

I also incorporate the realistic feature that high-level managers do not only exert productive effort on a given project. They also make investment and project selection decisions that affect the distribution of firm returns. In such a setting, managerial risk aversion becomes an extremely important consideration in contract design, since managers with different degrees of risk aversion make different investment decisions when faced with the same contract. The results provide new understanding of the roles played by incentive compensation, as well as the optimal degree of incentive power provided by such contracts.