Temporal Resolution of Uncertainty, the Investment Policy of Levered Firms and Corporate Debt Yields

Alexander S. Reisz*

October 26, 1999

(Job market paper)

---

* Doctoral Student, Finance Department, Stern School of Business, New York University, 44 W. 4th Street, MEC, Suite 9-190P, New York, NY 10012-1126. Tel: (212) 998 0344, Fax: (212) 995 4233, E-mail: areisz@stern.nyu.edu. I am grateful to my advisors, Douglas Gale, Kose John and Marti Subrahmanyam, for their guidance and numerous valuable ideas. I would also like to thank Yaakov Amihud, Stephen Brown, Eli Ofek and Gary Simon for their valuable suggestions, as well as Viral Acharya and Christopher Mann for their insightful comments. This paper would not have seen the light of day (at least not under its actual form) without the constant help and support from Kenneth Garbade. A special thought goes to Claudia Perlich and Jerold Weiss. The usual disclaimer applies.

Temporal Resolution of Uncertainty, the Investment Policy of Leveraged Firms and Corporate Debt Yields

Abstract

This paper attempts to link the agency literature (concerned with whether debt will trigger underinvestment incentives or risk-shifting behavior) with the one dealing with temporal resolution of uncertainty. To the best of our knowledge, apart from one article by John and Ronen (1990), there is no research article linking the two literatures.

We are concerned here with how the product/input market influences deviations from the optimal investment policy, in particular to what extent the speed of resolution of uncertainty of the industry in which a given firm operates affects the risk-shifting behavior of a shareholder-aligned manager. We assume that investors are risk neutral and that the return on the risky technology is normally distributed. It is then shown that the pattern of temporal resolution of uncertainty monotonically affects risk shifting as well as bond yields, even after contracts mitigating deviations from optimal investment policy have been written; empirical implications are derived and discussed.

1. Introduction

A recent review of the work done in the field of corporate finance over the last twenty five years ends with the conclusion that one of the biggest challenges facing theorists and empiricists alike is the valuation of "firms whose major asset consists of human rather than physical capital, prominent examples being in computer software and film production, […] for the paradigm example underlying most of our theoretical models is the manufacturing firm which dominated the growth of the economy around mid-century"[1]. This is a very large task, and we plan on tackling but one aspect of it: how the speed at which uncertainty is resolved for a firm/industry may affect its capital structure and the yield demanded on corporate debt.

The relevance of the concept of temporal resolution of uncertainty for corporate finance becomes intuitive as soon as one realizes that both sides of the balance sheet will depend heavily on it: the capital budgeting process will have to take into account whether more uncertainty affects short term or long term cash flows to figure out the right discount rates to use. It will also have to consider whether an expected value is a sufficient statistic to estimate future cash flows: if a discounted cash flow analysis seems appropriate for, say, a paper or lumber company, it is certainly not the case for a software or biotech firm, which will have to rely on other valuation techniques, for instance derived from option pricing theory. This, in turn, will determine the extent to which a firm can rely on debt and the type of debt the firm will use[2]. In this paper, however, we are not so much concerned about security design as about how the speed at which uncertainty is resolved will affect a firm's capital structure and the distortions between the optimal investment policy and the policy of a shareholder aligned manager.

The relation between capital structure and agency problems has been extensively studied. For the last thirty years[3], researchers in the field of finance have realized that leverage differs in a statistically significant manner across industries. Such a systematic and persistent variation across industries was surprising in light of Modigliani and Miller’s (1958) Capital Structure Irrelevance Theorem and much work was done to reconcile theory and evidence. Agency as well as signalling games were called on to explain the determinants of capital structure differences across industries and across firms within industries.

Agency theory focuses on problems where there is asymmetric information between managers/insiders and outsiders, in particular when managers can decide whether to undertake a risky project or a riskless one based on an experiment that they can privately observe. The central question in these papers is whether the presence of debt triggers risk-shifting (overinvesting in a risky technology or even investing in a negative net present

---

[1] Brennan (1995), p.18.

[2] See Goswami, Noe and Rebello (1995) for an attempt to design the type of debt a firm should use depending on whether more uncertainty surrounds short term or long term cash flows.

[3] Beginning with Schwartz and Aronson (1967).

value project) or underinvestment (Myers (1977) argued that if the risky debt outstanding matures after valuable investment opportunities, shareholder-aligned managers might consider that any positive result of an investment will go to bondholders and decide to strategically default instead of investing).

Two intuitions are at play here when one tries to figure out the effect of risky debt on the investment policy of a firm: on the one hand, considerations about the *nature* of the investment, or what I would call the "soccer" intuition[4]: a soccer team that is one or two goals behind will think only about attacking (possible positive returns) and be far more aggressive than is commonly deemed "reasonable" (NPV maximizing behavior); if it gives up one or two more goals, the situation is not much worse anyway: losing by two or four goals doesn't make a difference (for a shareholder, just defaulting or being badly in the red has the same effect: she loses her initial stake and nothing more due to the limited liability feature). Hence, appealing to the option pricing literature and realizing that the shareholders have sold the assets of the firm to the bondholders but have the option to buy it back for the face value of the debt ("strike price" of the option), it makes sense for them to maximize the value of this option by raising the volatility of the underlying assets[5]. Brito (1999), however, argues that the presence of (unrealized) growth opportunities known only to the manager will trigger an investment policy that is too conservative (not risky enough), since being driven into bankruptcy by lenders who do not observe those growth opportunities would simply mean losing them. As a consequence, growth-oriented firms might lean towards a lower debt/equity ratio so as to be able to invest freely in proprietary technologies. Finally, the manager may not be totally shareholder-aligned and may be concerned about the loss of control benefits or of reputation (especially if it is hard to tell whether a firm went bankrupt because of bad luck or bad management). This imperfect alignment between the manager and shareholders might lead her to be overly conservative in her investment policy and adopt a risk-avoiding behavior.

On the other hand, considerations arise about the *level* of investment expenditures. For instance, it has been argued that the manager of a corporation that issued debt will tend not to invest as much as would be optimal. This is due to the fact that returns have to be shared with bondholders, and that in the worst states of the world where the firm is still solvent, all of the returns go to bondholders (this is the debt underinvestment problem of Myers (1977) cited earlier). In other words, heavily levered firms may lower their level of physical capital investments in order to minimize the salvageable assets that would be lost to debtholders in the event of bankruptcy.

However, the agency literature has focused on very concrete industry characteristics (and understandably so, since it was the only way to obtain clear-cut empirical implications), and has left the effect of some concepts, such as temporal resolution of uncertainty (in the sequel: TRU), a yet… unresolved issue. We argue that if an

---

[4] The reader will understand that this paper was partly written during the soccer World Cup of 1998, and that the author didn't think of much else during that period.

[5] We remind the reader that the option price is increasing in the volatility of the underlying asset.

optimal capital structure or investment policy indeed exists, it will depend on the particular industry in which the firm operates, not only because of differences in expected bankruptcy costs or debt-related tax shields, but also because of the speed at which uncertainty is resolved in that particular industry. When making investment and employment decisions, a firm invests in specialized physical and human capital, which results in a cross-sectional variation in investment opportunity sets (i.e. prospective investment opportunities and associated payoff distributions). This should in turn help us understand cross-sectional variation in capital structure: the (private) release of new information might well affect cash flows due to agency games or adverse selection problems and "a simultaneous combination of optimal information generation and dissemination along with optimal financial contracts"[6] is to be determined: *the financial structure of a firm has to be designed based on the pattern of resolution of uncertainty this firm faces*. This explains why projects in their initial phase are often financed differently from projects in more advanced phases. It also sheds some new insight into why growth firms use a larger component of retained earnings for their financing than mature firms (which tend to generate more free cash flows), and why project financing (incorporated as legally segregated entities) is often used for the start-up phase of some ventures to switch later to more conventional modes.

The literature dealing with temporal resolution of uncertainty, however, is very sparse: the pioneering work by Epstein and Turnbull (1980) showed that when the uncertainty about a cash flow is resolved earlier (in the sense that more information is available about it at an intermediate date), the return demanded by investors who have a CARA utility function in an environment where returns are multivariate normal is higher than the riskless rate: the expected release of information has rendered the holdings risky. However, Epstein and Turnbull's discussion is restricted to studying the effect of temporal resolution of uncertainty on equilibrium market prices and on the optimal production of information. In their model, after the manager has conducted the experiment that yields some early resolution of uncertainty, "the firm communicates the [result of the experiment] truthfully" and "no production decisions are taken after the experiment results become known" (p. 628). They are aware that there might be "a moral hazard in that actions by management may not be in the best interest of the owners and will decrease the market value of their holding" and that the firm may release spurious information, but consider that insisting upon monitoring of managerial activities will minimize the associated costs. Hence, their conclusion that "the firm that maximizes market value will produce too little risk [insofar as] too little investment will be taken" (p. 638) has to be understood as totally independent of shareholders/bondholders conflicts.

Nabar, Stapleton and Subrahmanyam (1988) took the issue into the realm of corporate finance by studying how the value of corporate debt (and hence its yield) is affected when the speed of uncertainty resolution varies. However, their model assumes a context in which "the assets of the firm are in place and future risky investment decisions have already been taken by the firm" (p. 224). As a consequence, the Modigliani-Miller theorem still holds and "hence the leverage employed by the firm merely determines the split of total value

[6] John and Ronen (1990), p. 93.

between debt and equity and not the total value of the firm itself" (p. 225). Their model thus ignores agency problems and their effect on debt value and investment policy.

To the best of our knowledge, only two articles (John (1987) and John & Ronen (1990)) have tried to combine both literatures and examine agency problems under the light of temporal resolution of uncertainty. These two papers assume a discrete distribution for the outcome of the risky investment (a high outcome H and a low outcome L), look only at a few particular patterns of temporal uncertainty and thus can only offer a very limited idea of how temporal resolution of uncertainty affects risk-shifting incentives[7]; hence, the only conclusion they can draw is that "the financing choices and the investment policy can be materially affected by the timing of financing and investment vis à vis crucial resolutions of uncertainty in the underlying technology"[8]. We don't know of any research article trying to shift from such a qualitative statement to a quantitative comparative statics analysis, letting the pattern of temporal resolution of uncertainty be *arbitrary*.

Our goal here is to generalize the models discussed in the previous paragraph as well as to combine both the agency and the TRU literatures to analyze how different speeds of resolution of uncertainty influence the investment policy of the firm (i.e. risk-shifting or underinvestment) as well as the corporate debt yield. Our setup is different in that the managers don't observe the probability of success or the magnitude of final cash flows (as in most of the agency problems literature), but observe at t=1 the signal X_1, which gives them *partial* information about the return on the risky technology in which they can invest. This will lead us to various empirical implications, explaining, at least partially, why leverage and corporate bond yields differ consistently across industries.

The article is organized as follows: Section 2 posits the model and states the assumptions. Section 3 solves the problem when outsiders are not allowed to contract. It is shown in Section 4 that all our results carry over when bond covenants are written based on the outsiders observing a noisy signal. Section 5 derives empirical implications and confront them to the existing literature and Section 6 concludes. All proofs beyond the most trivial ones are relegated to the Appendix and numerical simulations appear in the end.

---

[7] In the same framework as the articles just quoted, John and Chidambaran (1998) combine an agency model with temporal resolution of uncertainty, but their model considers an all-equity firm and how much outside monitoring is needed to alleviate the problem of the manager investing suboptimally to maximize his compensation. Hence their model elaborates on the discrepancy between shareholders' and the manager's interests, and does not consider capital structure at all. On the other hand, we decided not to cut the umbilical cord that ties managers' acts to stockholders interests and to concentrate on the effect of a given pattern of temporal resolution of uncertainty on the debt ratio of the firm and the yield returned by this debt.

[8] John (1987), p. 638.

2. Model and Assumptions

We consider here a three-date (t=0,1 or2), two-period model[9]. The sequence of events is as follows:

- At t=0, the entrepreneur, who owns the rights to a firm but does not have enough capital to finance it, sells claims consisting of debt and equity to outside investors. The debt is sold entirely to outsiders, while the entrepreneur may retain some of the equity[10].
- At t=1, the manager of a firm with cash resources of I is faced with two possible investments: a riskless one, yielding the riskfree rate r_2-1 (known at t=1) and a risky project that yields the stochastic rate θ-1. She makes her investment decision based on the observation of a signal X_1. This signal is assumed to be bivariate normally distributed with the return available on the risky technology and hence gives some information to the manager about the probabilistic properties of the latter (we will denote ρ the correlation between X_1 and θ). Based on the observation of a particular realization x_1, the manager, a Bayesian decision maker, allocates her cash between the risky technology (in which she invests a fraction Q) and riskless Treasury bills in which she puts the remainder I-Q.
- At t=2, the realization of the risky technology θ is revealed to everybody and the firm is liquidated (proceeds are distributed according to the prewritten financial agreement).

The signal X_1 is not observed by outsiders, which precludes any contracting[11] (either managerial contracts or debt covenants) contingent on the value of X_1. This asymmetry of information creates a problem of incomplete contracting in relation to the entrepreneur's risk choices. Imperfect observability of private action (and resulting incomplete contracting) is at the heart of the agency problem. In Section 4, we will relax this assumption and allow outsiders to contract based on the observation of a noisy estimate of X_1.

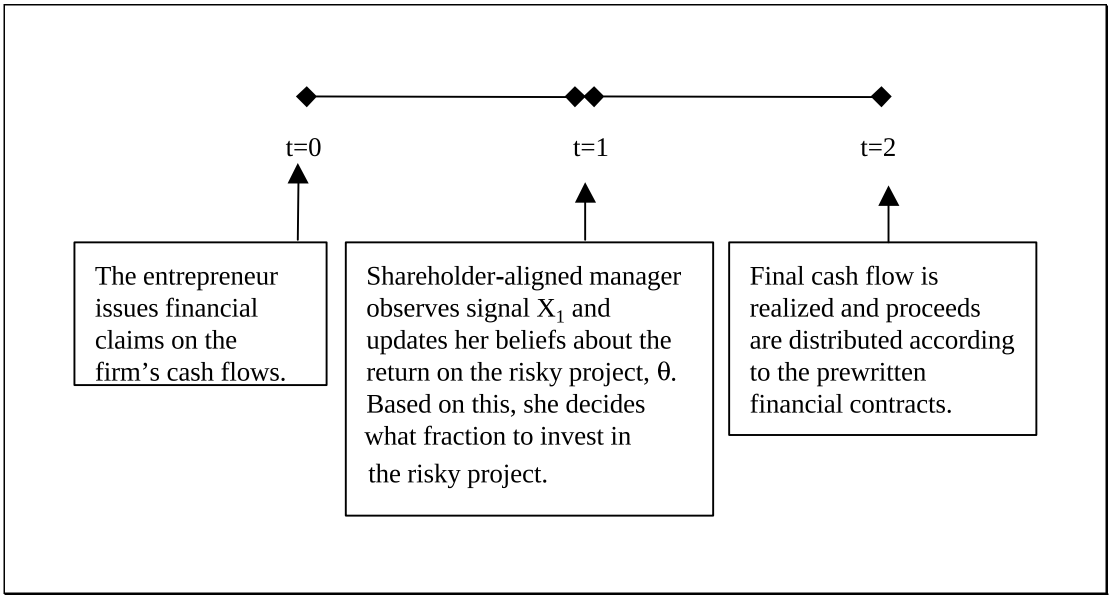
Our model can be summarized schematically as in figure 1:

---

[9] This discrete finite-period model is used not only to retain parsimony and elegance in a multi-period model, but also to have more flexibility in the correlation structure of cash flows: using a continuous time diffusion-type process would force us into a situation where we could not distinguish between the "pure" correlation element ρ and the time element $\sqrt{t/t'}$ that enters the covariance between two cash flows happening at different points in time t an t', t<t'.

[10] This insider contribution could be endogenized by trading off the reduction in contracting costs with the opportunity costs of supplying this capital, but we will abstain from it in this paper.

[11] X_1 could be the result of a performance test on a prototype or a marketing survey, accounting or market statistics, such as growth of sales or of net income, or a reflection of production efficiency; alternatively, it could be a t=1 cash flow, but we would have to impose the restriction that it is not contractible (or at least not verifiable by an outside court). We further assume that dividends are not allowed, in which case the shareholder-aligned manager would distribute as much as possible at t=1.



<u>Figure 1</u>

"Temporal resolution of uncertainty" is defined in much the same way as in Epstein (1980): we say that the more (less) informative the experiment concerning the final cash flows is, the earlier (later) the prior uncertainty about cash flows is resolved. Epstein and Turnbull (1980) show that in the case of jointly normal distributions, an experiment X_1 is said to be more informative about θ than about θ' if corr(X_1,θ)≥corr(X_1,θ'). In our model, we shall say that there is earlier resolution of uncertainty the larger (in absolute value) ρ, the correlation coefficient between X_1 and θ. This role of ρ is best seen if we consider the proportion of the uncertainty of θ that is resolved by observing X_1: $1-\mathrm{var}(\theta \mid X_1)/\mathrm{var}(\theta)=\rho^2$.

In other words, if ρ is high, there is little more to be learnt at t=2 and most of the uncertainty has indeed been resolved at t=1 (var(θ) is assumed to be a fixed parameter of the risky technology). We will consider only positive values of ρ, since we look at it as the correlation coefficient between θ and a random variable that is perfectly correlated to X_1 and which tells you how informative X_1 is about θ (for instance, a signal that has a -0.9 correlation with θ would be extremely informative; in our model, this case would be summarized with ρ=0.9).

We need here to clarify a few assumptions of our model:

- *<u>Assumption 1</u>*: Investors are risk-neutral. This stems either from the presence of complete markets[12], or from the combination of certain preferences and distribution properties for the underlying variables[13].

[12] To ensure the existence of a risk-neutral measure, we only need to assume away arbitrage.

As a consequence, the firm's securities are priced on the basis of the relevant conditional expectation of the cash flows (say, with respect to the state price density function).

- *Assumption 2*: The firm has two types of marketed claims outstanding, debt and equity. The debt has the form of a pure discount bond of promised payment F which matures at t=2. Note that since the return on the investment is normally distributed, it has a positive probability of being arbitrarily negative and therefore any amount of outstanding debt is risky. Once the final cash flow is realized, the firm is obliged to pay debt claims, if possible. If the cash flow is insufficient to meet debt obligations, the firm goes bankrupt and its assets are turned over to bondholders. We will abstain from considering bankruptcy costs or the tax advantage of debt financing.
- *Assumption 3*: It is not possible for the manager to issue further debt at t=1 after observing the result of the experiment, nor is it allowed to invest a negative amount in the risky project, putting the proceeds in the risk-free asset ("short-selling" the risky project). We also assume that $E(\theta)>r_2$, which we believe corresponds to the majority of situations.
- *Assumption 4*: The information available to the manager at t=1 is restricted to a signal X_1[14].
- *Assumption 5*: The manager acts to maximize the wealth of current shareholders (i.e., under asymmetric information, the true value of their claims conditional on the private information). In a rational expectations equilibrium, debtholders as well as stockholders will correctly anticipate, at t=0, the effect of debt structure and temporal resolution of uncertainty on the chosen risk strategy and the effect of this strategy on security pricing; in consequence, the entrepreneur bears the agency costs of debt when he sells securities at t=0. Note here that we implicitly assume that outsiders have all the information about the firm's characteristics (in particular about the firm's investment opportunity set) and insiders' preferences[15], which is necessary for them to rationally price their claims.

A few comments on those assumptions are necessary. Assumption 2 may not be as inocuous as it seems. In our model, the outcome of the risky project may be arbitrarily negative. Since in the United States (as in most

---

[13] Stapleton and Subrahmanyam (1984) show that in a discrete time framework, the combinations of negative exponential (CARA) utilities and normally distributed variables or power (CRRA or HARA) utilities and lognormally distributed variables yield a risk-neutral valuation relationship (RNVR) after a shift of the original parameters. See also Ross (1978) for why we do not lose any generality assuming risk neutrality if markets are "reasonably complete". In our model, the reader can consider that we look at the case where investors display constant absolute risk-aversion and returns are normal, but that all results are derived with parameters already shifted in such a way as to be able to use the tools of risk-neutral valuation.

[14] If one wants to relax this assumption, then all we need to assume is that if there were two different projects yielding θ and θ' respectively, all elements of the t=1 information set $\Im_1$ apart from X_1 are equally informative about θ and θ'. In other words, the manager considers only X_1 as a source of information for choosing how much she will invest in the risky technology.

[15] In other words, insiders' optimal private actions are common knowledge.

developed countries) equity has limited liability, we have to introduce a third claim. This claim is a purely negative one, borne by the government: if the firm faces hard times (i.e. a negative realization of θ), the government steps in to absorb the negative result (as, for instance, in the case of an environmental catastrophe where the state has to bear the cleanup costs)[16]. We will refer to the combined value of debt and equity as the "market value of the firm" since this is the amount for which the firm can be sold to the public at t=0. If we add to this the negative claim, we'll refer to it as the "social value of the firm".

Assumption 3 tells us that the firm chooses to issue outside claims (debt in our case) when it is common knowledge that there is no asymmetric information between the manager of the firm and outsiders. It is not clear whether it would be possible to issue further debt at t=1, when aymmetric information may lead to a "lemons" problem (see Akerlof (1970)), at an acceptable rate. This is an issue we are considering in a companion paper.

Finally, Assumption 5 clarifies what is known by everybody and what is private information: apart from the realization x_1 that only the manager observes, everything else is publicly known; in particular: ρ, the speed of resolution of uncertainty for the particular firm/industry, $E(\theta)$, $Var(\theta)$, $E(X_1)$, $Var(X_1)$ and all other parameters, as well as the preferences of the manager, the firm's characteristics and the fact that the manager privately observes x_1, thus enabling bondholders to rationally price their claim. All other assumptions are standard and do not affect crucially our results.

The problem here is to determine to what extent the industry in which the firm operates, or more precisely how quickly uncertainty is resolved in that particular industry, will affect the manager's deviation from the optimal investment policy and, as a result, the yield demanded on corporate debt.

3. Solving the Problem

3.1. Investment Policy of an All-Equity Firm

In this section, the investment policy of an all-equity firm with a managerial contract that perfectly aligns managerial incentives with shareholder interests is characterized. It defines the ideal benchmark, which would be obtained with complete conracting. At t=1, the manager seeks to

[16] Alternatively, this third claim could be an insurance policy for which the firm would have to pay a premium upfront. This complicates the model unduly, so we'll assume that the government steps in in case of a negative final cash flow but ensure in choosing our simulation parameters that the probability of this event happening is virtually zero.

$$\max_{Q\in[0,I]} \frac{1}{r_2} E\left[\left\{ Q\theta + (I-Q)r_2 \right\} | X_1 = x_1\right]$$

where $(w)^+$ stands for max(w,0). We will characterize this investment policy more fully once we have looked at the levered firm case, since the problem at hand can be seen as a particular case of a levered firm investment decision when the amount of outstanding debt F is 0. The analysis concerning the levered firm will then also apply to the case of the unlevered firm, replacing F by 0.

3.2. Investment Policy of a Levered Firm

Imagine now that the manager of the levered firm is shareholder-aligned. Her goal, when she decides at t=1 how much to invest in the risky technology based on the observed x_1, is now to:

$$\max_{Q\in[0,I]} \frac{1}{r_2} E\left[\left\{ Q\theta + (I-Q)r_2 - F \right\}^+ \mid X_1 = x_1\right].$$

Now the payoff to the shareholders is positive if and only if $Q\theta+(I-Q)r_2>F$ or

$$\theta \geq \frac{F-(I-Q)r_2}{Q} \equiv \theta^*$$

(note here the circularity of the problem: θ^* will determine the shareholders' optimal Q, but Q enters the formula for θ^*)[17]. The fact that θ^* increases with F reflects the fact that as the amount of promised debt increases, the range of "states" where the firm defaults, $(-\infty,\theta^*]$, expands[18].

The problem then becomes, denoting $P(\theta|X_1)$ the probability distribution function of θ conditional on X_1,

$$\max_{Q\in[0,I]} \int_{\theta^*}^{\infty} \left[Q\theta + (I-Q)r_2 - F\right] dP(\theta \mid X_1) \Leftrightarrow \max_{Q\in[0,I]} U(Q, r, x_1, F) \qquad (1)$$

---

[17] If the manager chooses to invest very little money in the risky technology (Q tends to 0), two possible cases arise: i) if $Ir_2 \geq F$, θ^* tends to $-\infty$: the firm will be solvent in all states of nature since its final wealth will be Ir_2, which is enough to cover its debt obligation of F; ii) if $Ir_2<F$, θ^* tends to $+\infty$: the firm will never be solvent since its final wealth, Ir_2, is insufficient to cover its debt obligation.

Now, it is hard to imagine that any bank or individual would lend money with a face value greater than Ir_2 to our firm (which, in order to be solvent, would *have to* invest a positive Q in the risky technology, no matter how negative x_1 is, triggering obvious risk shifting behavior; no rational bondholder would force shareholder-aligned managers, who, as we'll see shortly, already have a tendency to increase the volatility of the firm beyond what debtholders would find optimal, in a desperate situation where investing in the risky technology even if $E(\theta|X_1)<r_2$ is the only way to have a positive probability of avoiding bankruptcy). Therefore, in the sequel we shall assume that $F \leq Ir_2$.

[18] Unless otherwise stated, we call "state of the world" a particular x_1 if we are sitting at time 1, a particular θ or more generally a particular couple (x_1, θ) if we are sitting at t=2.

where we have underscored the dependence of this indirect utility function on ρ and the particular x_1 observed at t=1. It is shown in Appendix 2 that this is equivalent to maximizing the following expression over $Q\in[0,I]$:

$$U(Q,\rho,x_1,F)=\left[QE(\theta \mid X_1=x_1)+(I-Q)r_2-F\right]\Phi(B)+Q\sqrt{\mathrm{var}(\theta \mid X_1=x_1)}\,\varphi(B) \qquad (2)$$

where $B=\dfrac{E(\theta \mid X_1=x_1)-\dfrac{F-(I-Q)r_2}{Q}}{\sqrt{\mathrm{var}(\theta \mid X_1=x_1)}}$ and Φ and φ denote the standard normal cumulative probability function and density function respectively.

The expression (2) is the expected cash flows to the shareholders at t=2 and has to be discounted at the riskless rate r_2 to yield the share price at t=1, as a function of the quantity invested Q, the particular "state of the world" x_1 and the particular pattern of resolution of uncertainty (ρ shows up as well in $E(\theta|X_1=x_1)$ as in $\mathrm{var}(\theta|X_1=x_1)$).We can interpret this as follows: the first term tells us that the shareholders get whatever is left of the cash flows to the firm after bondholders have been repaid if the firm is solvent (which happens with a probability of Φ(B)); the second term tells us that shareholders should be concerned not only with whether they finish "in-the-money", but by how much (some people will call it a "convexity correction" factor[19]).

The manager will then maximize the function (2) over Q (given a certain pattern of temporal resolution of uncertainty ρ and after observing a particular x_1). In Appendix 3, the first and second derivatives with respect to Q are shown to be:

$$\frac{\partial U(Q,\rho,x_1,F)}{\partial Q}=\left[E(\theta \mid X_1=x_1)-r_2\right]\Phi(B)+\sqrt{\mathrm{var}(\theta \mid X_1=x_1)}\,\varphi(B) \qquad (3)$$

$$\text{and}\quad \frac{\partial^2 U(Q,\rho,x_1,F)}{\partial Q^2}=\varphi(B)\frac{(F-Ir_2)^2}{Q^3\sqrt{\mathrm{var}(\theta|X_1=x_1)}}>0 \qquad (4)$$

Since $\partial^2 U(Q,\rho,x_1,F)/\partial Q^2>0$, the quantity that maximizes the expression (2) is a corner solution. Whether Q=0 or Q=I will depend on the particular x_1 observed by the manager. This is due to the facts that the derivative (3) is increasing in Q and that $\lim\limits_{\substack{x_1\to-\infty\\ Q\to 0}}\dfrac{\partial U(Q,\rho,x_1,F)}{\partial Q}=-\infty$: for low values of x_1, investing everything in riskless bills may yield a higher value for the objective function.

Here is a quick synopsis of what $U(Q,\rho,x_1,F)$ will look like (these are sections of the response surface $U(Q,\rho,x_1,F)$ for various x_1's, but we realized that these cuts were more informative than showing the surface):

---

[19] Once the investment in the risky technology is made, the firm is long an asset that bears a normal return and thus has to cope with this feature of convexity. This is the same reason why, in the Black-Scholes (1973) option pricing formula, d_1 is bigger than d_2: it is not only about whether one ends up in the money, which happens with probability $\Phi(d_2)$, but also *by how much*.

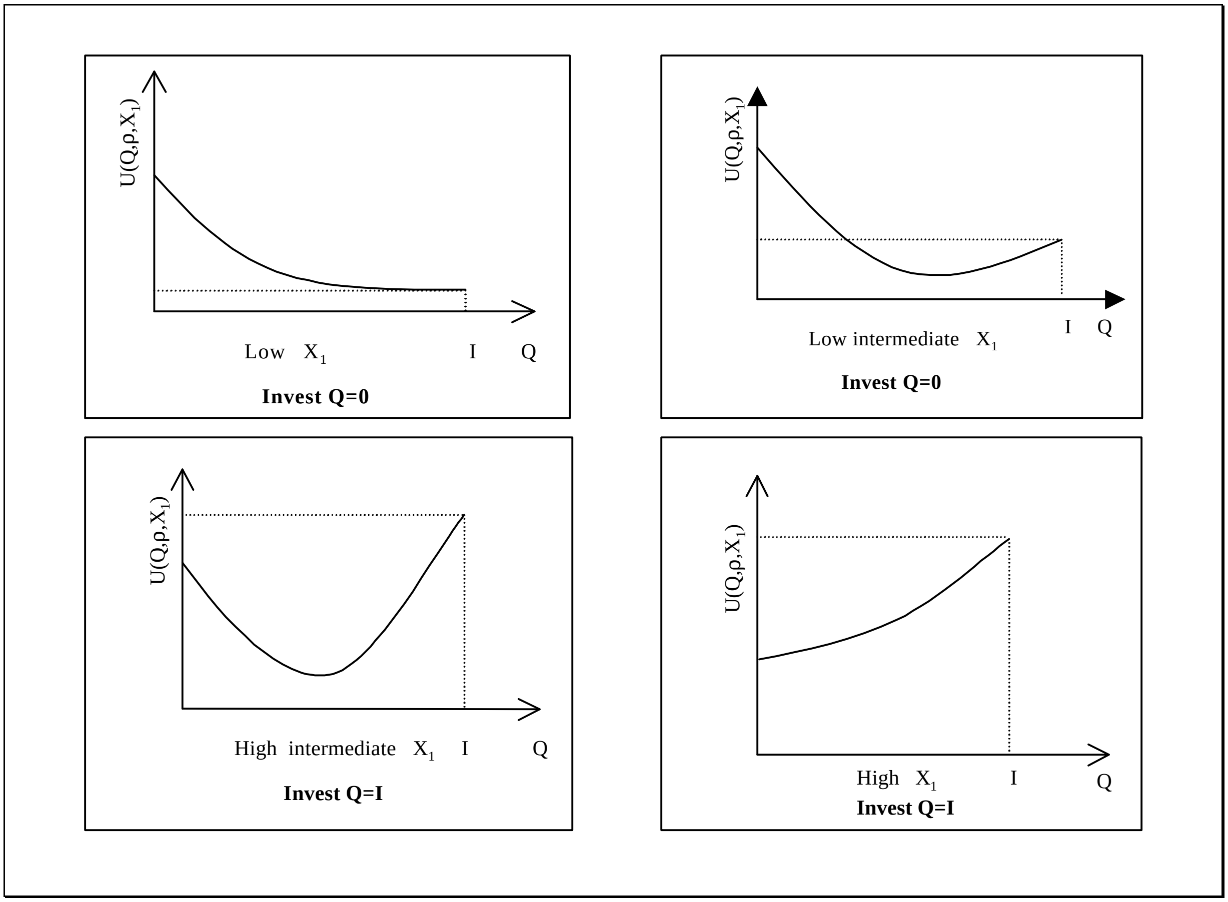


Figure 2

Our goal is to discover a range of x_1 values for which a full equity firm will not invest, but for which a levered firm will invest (i.e. a range of x_1 such that $U(I,\rho,x_1,F)>U(0,\rho,x_1,F)$ whereas $U(I,\rho,x_1,0)-U(0,\rho,x_1,0)\leq 0$). Then, we will be able to talk about full-fledged overinvestment (risk-shifting).

3.3. Characterizing the Risk-Shifting Incentives of a Shareholder-Aligned Manager

The first task is to find out whether there is an easy characterization of the x_1's that will trigger investment by the firm (i.e. a cutoff value X_1^F above which the firm will invest, below which the firm will not invest). It turns out we can answer this question positively.

Lemma 1: *Given a certain pattern of temporal resolution of uncertainty, there exists a unique cutoff value X_1^0 (resp. X_1^F) above which the manager of an all-equity(resp. levered) firm will invest in the risky technology.*

Proof: see Appendix 4.

The following definition facilitates the discussion and comparison of various investment policies:

Definition 1*: An investment policy of investing in the risky technology for all $x_1 \geq X_1^F$ will be denoted as investment policy $[X_1^F]$.*

Note that the investment policy $[X_1^0]$ is the one that could have been achieved if the realization x_1 was perfectly observed by all parties and if a complete set of enforceable contracts specifying any investment policy could have been written.

The next step is to show that a levered firm will never invest less that an all-equity one in the sense that whenever the all-equity firm will invest, so will the levered one[20]. This is the first important result of our paper:

Lemma 2*: The manager of a levered firm will have a strictly riskier investment policy than the manager of an all-equity firm in the sense that $X_1^F<X_1^0$. Hence, the investment policy $[X_1^F]$ gives rise to a cash flow distribution at t=2 that is riskier, in the sense of Rothschild and Stiglitz (1970, 1971), than the t=2 cash flow distribution from the investment policy $[X_1^0]$.*

Proof: see Appendix 5 for a proof that $X_1^F<X_1^0$. It then suffices to notice that over the range $[X_1^F,X_1^0]$, a strictly larger quantity will be invested in the risky technology by the levered firm, raising the variance of the final cash flows (while for any $x_1 \in [X_1^F,X_1^0]$, $E(\theta|X_1=x_1)<E(\theta|X_1=X_1^0)$).

Corollary 1*: The ex-ante probability of investing in the risky technology is strictly greater for a levered firm than for an all-equity one.*

Proof: this follows trivially from Lemma 2 and the normality of X_1: the probability that a firm with an investment policy $[\xi]$ will invest is $1-\Phi((x-\overline{X}_1)/s_{X_1})$; the fact that $X_1^F<X_1^0$ yields the corollary.

We therefore decide to call the quantity X_1^0-X_1^F the extent of risk-shifting, since it tells us how bad a signal x_1 can be relative to the cutoff X_1^0 that an all-equity firm will use while still seeing the manager of the levered firm investing in the risky technology. The following figure will help us clarify the situation:

---

[20] In other words, the X_1-range that triggers investment by the levered firm will contain the one that triggers investment by the full equity firm.

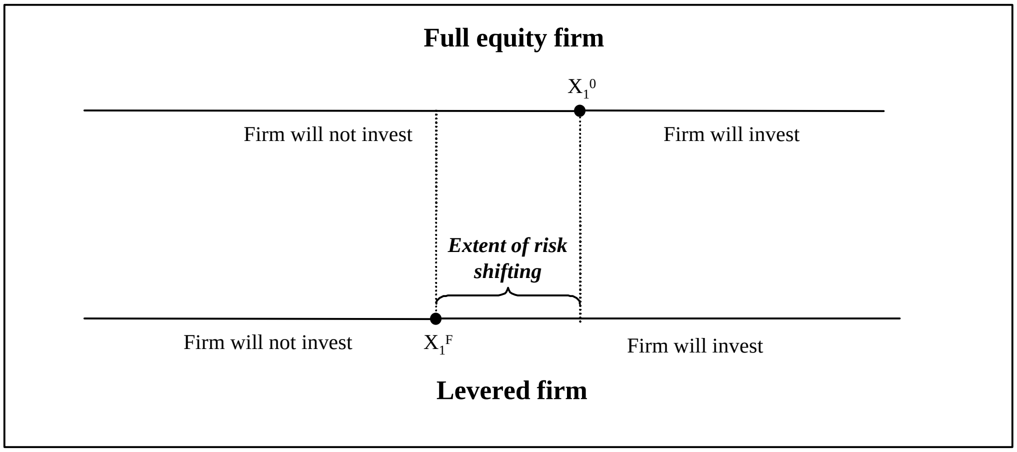


Figure 3

The obvious question that arises is about the relation between the extent of risk-shifting and the pattern of temporal resolution of uncertainty. This will be analyzed shortly and yield one of the central results of this study. But before that, a last preliminary result has to be stated:

Lemma 3: *For a given* r*, the extent of risk shifting is strictly increasing in the face value F of the firm's debt. Hence, given lemma 2, the terminal cash flow distribution resulting from the investment policy* $[X_1^F]$ *implemented by the manager is strictly increasing in risk for increasing F.*

Proof: in order to prove Lemma 2, we showed in the appendix that X_1^F is decreasing in F. Since X_1^0 doesn't depend on F, the quantity X_1^0-X_1^F, which was previously shown to be positive for any positive F, is therefore increasing in F. The lower cutoff value X_1^F resulting from an increase in the amount of outstanding debt yields the added riskiness. QED.

This corresponds to the intuition that the higher the probability of insolvency, the higher the incentive to increase the risk of the firm. It is also consistent with John and John (1993) and Brito (1999), as well as with the early intuitition of Jensen and Meckling (1976).

3.4. Risk-Shifting Incentives and the Pattern of Temporal Resolution of Uncertainty

We have just seen that the higher the amount of risky debt outstanding, the more pronouced the distortion in the investment policy of a shareholder-aligned manager. This is hardly a new result, although it is interesting to realize that it still holds in our model. However, a far more interesting task is to find out how risk shifting incentives vary with the pattern of temporal resolution of uncertainty. Several results will be proven here.

***Theorem 1**: (a) The cutoffs X_1^0 and X_1^F are strictly increasing in r.*

(b) As a result, the riskiness of the the firm's t=2 cash flow distribution is decreasing in r.

Proof: see Appendix 6 for a proof of (a). As for (b), it follows directly from Lemma 2: if a higher ρ triggers a higher cutoff value X_1^0 (resp. X_1^F), then the terminal cash flow distribution resulting from the investment policy $[X_1^0]$ (resp. $[X_1^F]$) implemented by the manager is strictly less risky.

This yields a directly testable implication we will talk about later: firms operating in industries where uncertainty is resolved only late have more risky operations. Heuristically, firms with later temporal resolution of uncertainty (lower ρ) will typically rely more on growth projects, developing a new drug or software, with the risk that a competitor achieves it before them or that a regulatory agency prevents them from selling their product; those projects are inherently more risky than the operations of well established firms operating in a field where the outcome of operations are quite forecastable. But our model says a little more: the added riskiness comes not only from the *nature* (σ_θ) of the firm's operations, but also from the speed at which the uncertainty of its operations is resolved: when the intrinsic riskiness of a project is held constant at σ_θ (i.e. even when we compare two firms in the same industry), the firm for which uncertainty is resolved later will invest over a larger X_1-region than the firm for which uncertainty is resolved earlier. In other words, if a given firm decides to invest for a certain value of x_1, so will a firm facing more delayed uncertainty[21]. This is due to the fact that when ρ is small, X_1 is less reliable in forecasting θ; hence the manager is ready to take the risk of investing in the risky project even for relatively low values of x_1. On the contrary, when ρ is larger, X_1 is more reliable in forecasting θ and she is not ready to take as much risk in her investment policy if the state of the world as of t=1, x_1, is mediocre; hence she sets her cutoff higher. Our results apply therefore primarily in the comparison of two otherwise identical firms (i.e. facing the same investment opportunity set, in particular a same σ_θ) but facing different patterns of temporal resolution of uncertainty. Of course if we let σ_θ vary and be greater for firms with more delayed resolution of uncertainty (which corresponds to our intuition), our results would only be reinforced[22].

The previous result holds for all-equity as well as levered firms. However, it seems a far more interesting question to find out how the extent of debt-induced risk shifting (the difference between X_1^0 and X_1^F, both of which increase with ρ) depends on how quickly the uncertainty is resolved, or in other words to talk about *the interaction of the product market and the capital structure of the firm*. This is done in the following theorem:

---

[21] Alternatively, a firm facing more delayed resolution of uncertainty invests more *often* in a probabilistic sense since $1-\Phi((x-\overline{X}_1)/s_{X_1})$, the probability of investing, is decreasing in ρ for $\xi \le X_1^0$.

[22] It is easily proved that $\partial X_1^0/\partial\sigma_\theta<0$ and $\partial X_1^F/\partial\sigma_\theta<0$ (proof available from the author on request).

Theorem 2: *The extent of risk shifting X_1^0-X_1^F is strictly decreasing in the informativeness* r *of the experiment.*

Proof: see Appendix 7.

The intuition is the following: the manager of a firm operating in an industry where uncertainty is resolved only late will find it a simple matter to deviate significantly from the optimal investment policy; the results of such a deviation would take years (in a model allowing for the uncertainty to linger beyond t=2) to be noticed by creditors, by which time the value of the firm might be eroded beyond repair. Potential creditors will shy away from a situation they could neither monitor nor control and we would expect the level of debt in such an industry to be lower than in industries where uncertainty is resolved earlier (where the extent of risk-shifting is not only lower as we have just proved, but also more effectively monitorable)[23]. As a matter of fact, as we will see later, this is supported by empirical evidence.

Note that our results are consistent with the existing research predicting that "companies whose value consists primarily of "growth options" as opposed to "assets in place" are likely to find debt financing very costly" (Myers (1977), p. 161). Intuitively "assets in place" are more numerous in industries where uncertainty is resolved earlier. Our paper, however, is the first one to uncover the functional dependence between the pattern of temporal resolution of uncertainty and both the overall risk of the firm and the investment policy distortions. Although our conclusions apply primarily to firms facing the same investment opportunity set but differing only through the speed at which they expect uncertainty to be resolved, a generalization to a cross-industries conclusion is straightforward[24].

Our conclusion about how firms with delayed uncertainty will find debt financing costly is also in line with DeAngelo and Masulis (1980): tax shields due to the deductibility of debt interest and non debt-related tax shields are substitutes and hence should be traded off against each other. R&D and advertising expenses (generally considered a good proxy for how late uncertainty is resolved) are investments that provide a greater tax shield than capital spending, because the entire outlay can be expensed for tax purposes in the year

---

[23] This is reinforced by Theorem 1, which states that the lower ρ is, the higher the variance of the t=2 cash flow distribution, decreasing in turn the debt capacity of the firm.

[24] It is easily shown that the extent of risk-shifting is increasing in σ_θ (proof available from the author upon request): a shareholder aligned manager who knows that the risky technology has become riskier considers only the fact that the probability of default has increased and will "go for broke"; in other words, she has an incentive to increase the risk of the firm to maximize the return in the states of the world where the firm is solvent. Therefore the "temporal resolution of uncertainty effect" would only be reinforced by the "risk effect" given that firms operating in fields where uncertainty is resolved only later also typically face a riskier investment opportunity set (higher σ_θ).

incurred instead of being amortized over time. Hence, they effectively diminish the benefit of the interest deductibility of debt because they reduce the number of states of the world in which there is unsheltered income remaining after these deductions are considered. One would therefore expect firms with high levels of R&D and advertising expenses to rely less heavily on debt[25].

It can also be shown that the extent of risk shifting is convex in ρ. It therefore goes very quickly to zero and in the limit, when ρ tends to 1, there is no risk shifting anymore: observing x_1 fully reveals θ and whether the firm is levered or not, the risky project will be chosen if and only if $\theta(x_1) = \bar{\theta} + \sigma_\theta \left(x_1 - \bar{X}_1\right) / \sigma_{X_1} \geq r_2$[26]. However, this feature of continuity does not hold when ρ tends to 0: when ρ=0, any particular x_1 is irrelevant in forecasting θ and the best estimate of θ is the unconditional mean, E(θ). Hence, following the same argument as in the footnote 26, both managers will invest under the same condition, i.e if $E(\theta) \geq r_2$ regardless of the particular realization x_1 and the extent of risk shifting is zero. However, when ρ tends to 0, we show in Appendix 7 that the extent of risk shifting tends to infinity. Hence, there is a discontinuity in the extent of risk shifting in ρ=0, but as we will see later, this does not create a discontinuity in the agency costs of debt in the neighborhood of ρ=0[27].

---

[25] There is an obvious argument against this: older, more mature firms will have higher levels of capital investment and hence higher depreciation (non debt-related tax shields), which is a substitute for debt-related tax shields. However, we believe that the fact that R&D expenses can be totally expensed instead of amortized will dominate the fact that R&D intensive firms will have lower levels of depreciation. Another element that supports our theory is that investments in firm-specific human capital can also be 100% expensed in the year they are incurred, and these non-debt tax shields are of course expected to be higher in a firm with late temporal resolution of uncertainty. See our section on empirical implications and evidence.

[26] When θ is revealed as of t=1, it is obvious that the manager of the full equity firm invests in the risky project if and only if $\theta(x_1) \geq r_2$ where $\theta(x_1)$ is the θ revealed by the realization of the particular x_1. However, the manager of the levered firm wants to maximize $[Q\theta(x_1)+(I-Q)r_2-F]^+$. If the expression in the square brackets is positive, then she will also seek to maximize $Q\theta(x_1)+(I-Q)r_2$, i.e. invest in the risky project if and only if $\theta(x_1) \geq r_2$; any lower realization x_1, and the manager is better off investing in riskless bonds, securing a payoff of Ir_2-F for shareholders.

[27] The fact that the extent of risk shifting tends to infinity when ρ approaches 0 does not have any impact on agency costs for two reasons: i) both X_1^0 and X_1^F tend to $-\infty$ (when $E(\theta) > r_2$), and for all relevant purposes, we can consider that both managers have the same investment policy (investing for all x_1's): the distortion in investment policy triggered by the presence of risky debt is negligible (in this most extreme case, it might be of more relevance to consider the *probability* of falling in this region, $\Phi((X_1^0 - \bar{X})/\sigma_{X_1}) - \Phi((X_1^F - \bar{X})/\sigma_{X_1})$, which indeed tends to 0); ii) when ρ is very low, X_1 is of very little relevance for predicting θ and even a significant distortion between the two investment policies would not yield significant agency costs.

3.5. Firm Value, Agency Costs and Temporal Resolution of Uncertainty

The fact that the extent of risk shifting X_1^0-X_1^F is strictly decreasing in the informativeness of the experiment is not enough *per se* to come to the conclusion that agency costs are also decreasing in ρ. As a matter of fact, it is not true. However, it is still of interest to briefly study the agency costs of debt in our model.

We show in Appendix 8 that the t=0 *market* value (i.e. the value of both debt and equity) of a firm with an investment policy of [ξ] (i.e. [X_1^0] for the all-equity firm, [X_1^F] for the levered one) is equal to:

$$V_0(\rho,\xi)=\frac{I}{r_1r_2}\left\{r_2\Phi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)+\sigma_\theta(1-\rho^2)^{1/2}\int_\xi^\infty\left[A\Phi(A)+\varphi(A)\right]dP(X_1)\right\} \qquad (5)$$

where A stands for $E(\theta\,|\,X_1)/\sigma_{\theta|X_1}$ and is a function of X_1 and r_1 denotes the discount rate to be used from t=0 to t=1 (we would not lose any generality assuming that it is equal to 0)[28].

It is interesting to note that the value of both the all-equity and the levered firm is increasing in ρ (see Appendix 8). We can see $\partial V_0(\rho,\xi)/\partial\rho>0$ as the marginal increase in firm value that a marginally more "educated" choice by the manager will bring (marginal value of information) [29]. If the model allowed the manager to buy more information given a certain cost, this increase in marginal value would have to be

[28] Note that the above is different from the *social* value of the firm, which is equal to:

$$V_0^S(\rho,\xi)=\frac{I}{r_1r_2}\left\{\bar{\theta}-(\bar{\theta}-r_2)\Phi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)+\rho\sigma_\theta\varphi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)\right\} \qquad (6)$$

NB: This social value can be decomposed in three terms: the first one, $IE(\theta)/(r_1r_2)$, tells us that as of t=0, the firm's cash is expected to appreciate at the unconditionally expected rate on the risky technology, E(θ); the second one, $I(r_2-\bar{\theta})\Phi((\xi-\bar{X}_1)/\sigma_X)/(r_1r_2)$, can be seen as a "negative risk premium": with probability $\Phi((\xi-\bar{X}_1)/\sigma_X)$, i.e. for all $x_1<\xi$, the firm will *not* invest in the risky technology and will therefore, on average, earn r_2 instead of E(θ) on its cash, losing the risk premium $\bar{\theta}-r_2$; finally the third term represent a kind of "convexity adjustment".

The difference is accounted by the third negative claim, i.e. the responsibility to cover a negative realization of θ (e.g. environmental clean up costs) which we denote $G_0(\rho,\xi)$:

$$G_0(\rho,\xi)=\frac{I}{r_1r_2}\left\{\bar{\theta}\left[1-\Phi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)\right]+\rho\sigma_\theta\varphi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)-\sigma_\theta(1-\rho^2)^{1/2}\int_\xi^\infty\left[A\Phi(A)+\varphi(A)\right]dP(X_1)\right\}$$

[29] An unfortunate feature of our model is that ρ is as well the pattern of resolution of uncertainty as the extent of asymmetric information the manager has at t=1 (how reliable the x_1 observed only by her is). Considering this, we might expect an increase in the amount of asymmetric information ρ to trigger an increase in the extent of risk-shifting and a decrease in firm value as is the case in the existing literature; however we find the contrary: *the temporal resolution of uncertainty effect overwhelms any opposite asymmetric information effect, and disentangling the two roles of ρ would only make our results stronger*. This is investigated in a later section of this paper.

equated to the marginal cost of buying more information to find the optimal level of information[30]. Note finally that the increase in firm value due to an increase in ρ does not only happen through a less significant deviation from optimal investment policy, since $V_0(r, X_1^0)$ is also increasing in ρ: the increase in firm value comes as well from a better investment policy (lower deviation $X_1^0 - X_1^F$) ("reduction in agency games effect") as from a lower residual uncertainty facing the manager: $\text{var}(\theta|X_1)=\sigma_\theta^2(1-\rho^2)$ ("total firm value effect"). Given the increasingness of firm value in ρ, it is therefore bounded below by $V_0(0, X_1^0) = V_0(0, X_1^F)$, both of which are greater than $\text{IE}(\theta)/r_1r_2$[31].

Since in our model $E(\theta)>r_2$, this lower bound for the firm's value is strictly greater than I/r_1, the t=0 present value of the I dollars invested at t=1. The difference can be seen as the "growth options" of the firm (value of the opportunity to invest in a technology with $E(\theta)>r_2$). Alternatively, it can be seen as the surplus of the entrepreneur when he sells the firm to outside shareholders and bondholders at t=0.

The agency costs of debt in our framework are thus equal to:

$$V_0(r, X_1^0) - V_0(r, X_1^F) = \frac{I}{r_1 r_2}\left\{ r_2\left[\Phi\left(\frac{X_1^0 - \overline{X}}{s_{X_1}}\right) - \Phi\left(\frac{X_1^F - \overline{X}}{s_{X_1}}\right)\right] + s_q(1-r^2)^{1/2}\int_{X_1^0}^{X_1^F}\left[A\Phi(A) + j(A)\right]dP(X_1)\right\}$$

We show in Appendix 8 that these agency costs are strictly increasing in the amount of outstanding risky debt F, which is in line with most of the existing literature but in contradiction with the earlier work of Gavish and Kalay (1983).

What is more interesting is that we get, in our numerical simulations, a bell-type curve, with very low agency costs for low and high values of ρ. Overall, agency costs are not very large: for I=100, r_2=5%, $E(\theta)$=15%, and σ_θ=0.6, agency costs represent up to 0.23% of firm value for F=50 (market value of debt/equity ratio: 0.675) and 1.16% for F=75 (market value of D/E ratio: 1.375). This is consistent with Leland (1998), who finds agency costs ranging from 0.32% to 1.22% of firm value depending on the risk management policy of the firm. The explanation for such low agency costs (especially for extreme values of ρ) is the following:

- When ρ is low, the extent of risk shifting X_1^0-X_1^F might be high, but the realization of θ given a low x_1 might still be decent since the signal X_1 is of so poor forecasting ability. Hence, agency

[30] That is, if her goal was to maximize firm value. If she is shareholder-aligned, as in our model, she would only be ready to pay the marginal increase in *equity value* due to an increase in ρ. It can indeed be shown that share value (as well as the value of the claim unwillingly held by society at large) is also increasing in ρ.

[31] When ρ=0, there is no discrepancy between X_1^0 and X_1^F: X_1 gives no information about θ, both managers compare the unconditional $E(\theta)$ to r_2 and always invest. The market value of the firm in that case is $\text{IE}(\theta|\theta\geq 0)/r_1r_2$ = $I\left[\overline{q}\,\Phi(\overline{q}/s_q) + s_q\, j(\overline{q}/s_q)\right]/r_1r_2 > I\overline{q}/r_1r_2$, this last number being the social value of the firm for ρ=0.

costs do not suffer too much from a low ρ (mathematically, $E(\theta|x_1) = \bar{\theta}+\rho\sigma_\theta(x_1- \bar{X}_1)/\sigma_X$, and even if x_1-$E(X_1)$ is very negative, $E(\theta|X_1=x_1)$ will be close to $E(\theta)$ for very low ρ). Another way of looking at it is to say that bondholders know at t=0 that managers won't have a lot of information at t=1 (ρ is public knowledge since it summarizes either the investment opportunity set of the firm or the industry in which the firm operates) and hence are not too worried about the consequences of asymmetric information[32]; in the limit, when ρ=0, X_1^0 and X_1^F tend to $-\infty$ (since X_1 does not give any information on θ, the manager compares the unconditional expectation of θ with r_2 and always invests since $\bar{\theta}>r_2$), agency costs are equal to 0 and the firm market value is equal to $IE(\theta|\theta \geq 0)/r_1r_2$, a lower bound for the market value of the firm, as seen previously.

- When ρ is very high, a very negative x_1 will be a good hint that θ will be low, but as we saw earlier, the extent of risk-shifting X_1^0-X_1^F will be very low, so that the manager of a levered firm will have an investment policy very similar to the one of an all-equity manager. Another way of looking at it is to say that bondholders are not worried about extensive asymmetric information as they know that the manager will have a policy very similar to the one they would undertake if they were in charge; moreover, the fact that the manager knows so much limits the residual risk of the firm. In the limit, when ρ=1,

$$V_0(1) = V_0^s(1) = V_0^F(1) = \frac{I}{r_1r_2}\left[\bar{\theta} + (r_2-\bar{\theta})\Phi\left(\frac{r_2-\bar{\theta}}{\sigma_\theta}\right) + \sigma_\theta\,\varphi\left(\frac{r_2-\bar{\theta}}{\sigma_\theta}\right)\right]$$

  The intuition is the same as before, the probability, *ex ante*, that x_1 will fall short of $X_1^0=X_1^F$ tending to $\Phi((r_2-\bar{\theta})/\sigma_\theta)$.

- The highest agency costs are achieved for intermediate values of ρ: if the extent of risk-shifting is smaller than for lower values of ρ, it also causes much more harm since a mediocre x_1 is a good hint that the final realization of θ will indeed be poor. The reason why the magnitude of agency costs in this region is still modest is that managers do not observe, as opposed to previous studies, any future parameter (e.g. future cash flows in case of success), but a signal that give them some *partial* information about future cash flows. Hence i) the magnitude of the asymmetric information is much smaller, and ii) if the policy conducted by the manager is much different in terms of X_1-cutoff from the one value-maximizers would choose, it does not ensure in any way that the final outcome of θ will make this deviation of consequence. Our agency costs are also rendered smaller by the fact that we do not consider the costs of financial distress (not only bankruptcy costs but also the impossibility to react freely to strategic moves by competitors if the firm is in the process of reorganization).

[32] As noted before, ρ also measures the extent of asymmetric information.

Now, the fact that agency costs are always positive would hint at an easy solution to the agency problem: regardless of the pattern of temporal resolution of uncertainty, firms should be all equity financed, thereby avoiding suboptimal managerial behavior. This reasoning is wrong given that we have abstracted, in our model, from considering the different benefits of debt (tax deductibility of interest, disciplinary and signalling role of debt) as well as agency costs of equity. *However, these benefits cannot be traded off against the agency costs of debt independently of the pattern of temporal resolution of uncertainty, upon which the optimal capital structure of the firm will depend*. For instance, the tax bracket of a given firm certainly depends on the investment set of that firm, i.e. on the industry in which it operates. Moreover, the dollar amount for which insiders can sell at t=0 the promise to repay F dollars at t=2 if the firm is solvent depends on the speed at which the uncertainty is resolved for the firm.

3.6. Temporal Resolution of Uncertainty and Corporate Debt Yield

We have shown that the earlier the resolution of uncertainty, the lower the overall riskiness of the firm (Theorem 1) and the extent of risk-shifting (Theorem 2). We would therefore expect rational bondholders to anticipate this and to pay a lower price (or, equivalently, demand a higher risk premium) for the bonds of a firm operating in a field where uncertainty is resolved later. This is actually something that we prove in Appendix 9, after computing the t=0 value of bonds:

$$B_0(\rho, F, X_1^F) = \frac{1}{r_1 r_2}\left\{F\Phi\left(\frac{X_1^F - \overline{X}}{\sigma_{X_1}}\right) + I\sigma_\theta (1-\rho^2)^{1/2}\int_{X_1^F}^{\infty}\left[A\Phi(A) + \varphi(A) - C\Phi(C) - \varphi(C)\right]dP(X_1)\right\}$$

Theorem 3*: The equilibrium prices of corporate bonds are increasing in ρ, the pattern of temporal resolution of uncertainty. Equivalently, the default premium demanded on corporate bonds is decreasing in ρ.*

Note that this stems from two effects:

i) a higher ρ increases overall firm value (as seen earlier) through the fact that the manager can carry out a more "educated" investment policy (alternatively, the residual variance of the risky project, if it is entered into, is lower for a higher value of ρ: $var(\theta|X_1)=\sigma_\theta^2(1-\rho^2)$); bondholders share this benefit with shareholders ("total firm value effect");

ii) as ρ increases, the extent of risk-shifting X_1^0-X_1^F decreases, and bondholders benefit from a lower deviation from socially optimal investment policy ("reduction in agency games effect").

Now, this result was to be expected: as shown in Theorem 1, the riskiness of the firm's t=2 distribution of cash flows is decreasing in ρ. Hence the decreasingness of the yield premium in ρ could be due solely to the fact that as ρ increases, the probability of bankruptcy decreases. We however bring something new to the old wisdom that the riskier a firm is, the higher the yield demanded by bondholders: in our model, *both the added riskiness of the firm's cash flows and the added extent of risk shifting (the other reason for higher yields) are due to the uncertainty being resolved later*. In a companion empirical paper (Reisz (1999)), we investigate whether temporal resolution of uncertainty still has some power explaining the cross-sectional variation in yields demanded on bonds, *once risk has been controlled for*. The results of this empirical investigation leads us to answer positively[33], but more empirical work is warranted at this point to confirm our simulations: using the same parameter values as before, the yield premium is as high as 383 basis points for F=50 (market debt/equity ratio of 0.675) and 628 basis points for F=75 (market D/E ratio of 1.375). This is consistent with Leland (1998) for instance.

We also prove in Appendix 9 that as ρ tends to 1, bond prices tend to the price of a risk-free bond (the default yield premium tends to 0). The intuition is straightforward: when ρ is equal to 1, θ is revealed as of t=1 and the manager invests in the risky project if and only if $\theta(x_1)>r_2$ (see footnote 26). Since $Ir_2>F$, the firm will always be solvent.

It is finally worth noting that, as was the case for firm value, *the temporal resolution of uncertainty effect overwhelms any opposite asymmetric information effect*: if ρ represented only the amount of information asymmetry between insiders and outsiders, we would expect rational bondholders to demand a higher yield for a higher ρ, as well as shareholder-aligned managers to risk-shift over a wider X_1-region. The fact that we find the opposite result tells us that the latter effect is dominated by the temporal resolution of uncertainty effect, and outsiders are willing to accept a higher level of information asymmetry if that means that the residual risk of the firm is lower. However, it would be nice to somehow disentangle the two effects of ρ and allow outsiders to contract (at least partially) so as to mitigate the extent of risk-shifting. This is done in the next section.

4. Mitigating Risk-Shifting with Partial Contracting

4.1. The Framework

So far, we considered that insiders had information as of t=1 about the t=2 outcome of the risky technology, but that outsiders had no information whatsoever. A more general model would allow for outsiders to also

[33] We also find in our empirical work that firms with more delayed resolution of uncertainty issue shorter bonds This was to be expected, since firms operating in a low ρ environment may want to keep the flexibility of refinancing under much better terms (lower interest rate) at t=1 if no investment is made. For firms with earlier resolution of uncertainty, this is less of an issue since their default premium is quite low.

have some information as of t=1, albeit less precise than the insiders' signal. This is done by letting the outsiders observe at t=1 a signal Y_1, which is a noisy version of the signal X_1 observed by managers:

$$Y_1 = X_1 + \varepsilon \quad \text{where} \quad \varepsilon \sim N(0, \sigma_\varepsilon^2) \text{ is white noise, independent of } X_1 \text{ and } \theta^{34}.$$

Proposition 1: *Y_1 and θ are bivariate normal random variables with correlation coefficient $\rho_b = \rho_m \rho_{XY}$ where ρ_{XY} is the correlation coefficient between X_1 and Y_1 and ρ_m is the correlation coefficient between X_1 and θ (informativeness of manager's signal).*

Proof: $\rho_b = \dfrac{\text{cov}(\theta, Y_1)}{\sigma_\theta \sigma_{Y_1}} = \dfrac{\text{cov}(\theta, X_1 + \varepsilon)}{\sigma_\theta \sigma_{Y_1}} = \dfrac{\sigma_\theta \sigma_{X_1} \rho_m}{\sigma_\theta \sigma_{Y_1}} = \rho_m \dfrac{\sigma_{X_1}}{\sigma_{Y_1}}$. Now $\rho_{XY} = \dfrac{\text{cov}(X_1, X_1 + \varepsilon)}{\sigma_{X_1} \sigma_{Y_1}} = \dfrac{\sigma_{X_1}}{\sigma_{Y_1}}$.

Hence $\rho_b = \rho_m \rho_{XY}$. QED.

Given that $\text{var}(Y_1 | X_1) = \sigma_\varepsilon^2 \equiv \sigma_{Y_1}^2 (1 - \rho_{XY}^2)$, it becomes clear that Y_1 is a more noisy estimate of X_1 (larger σ_ε) if and only if ρ_{XY} is smaller. It becomes in turn a less reliable predictor of θ (smaller ρ_b) and leaves outsiders with a higher residual uncertainty $\sigma_{\theta|Y_1}^2 = \sigma_\theta^2 (1 - \rho_b^2)$. Hence, the larger $\rho_m^2 - \rho_b^2$ and σ_ε^2 or the smaller ρ_{XY}, the larger the extent of asymmetric information.

A useful way to think about it is as follows: ρ_m is the informativeness of the manager's signal and is the true speed of temporal resolution of uncertainty in a given industry; it affects, as we have seen before, the extent of risk-shifting, firm value and corporate bond yields. It therefore gives us *inter-industry* comparative statics (once we have controlled for the intrinsic project risk σ_θ). However, *within* a given industry, disclosure requirements vary (due to size differences or the stock exchange on which a stock is traded) and therefore outsiders will be more or less able to monitor actions by insiders. Typically, large firms listed on the NYSE will have stricter disclosure requirements (higher ρ_{XY} or equivalently lower σ_ε) than smaller firms listed on the NASDAQ, leaving outsiders with a higher "signal precision" ρ_b[35]. This ρ_b *cannot*, however, be seen as any

[34] ε is assumed to be added by "nature" and not by the insiders. They may have some control over its distribution (i.e. σ_ε), but they cannot pick a certain realization ε_1 to add to x_1. For such a model in the context of "earnings management", see Degeorge, Patel and Zeckhauser (1999).

[35] ρ_b thus measures the speed of resolution of uncertainty "tainted" by asymmetric information. Note also that when we talk about "disclosure requirements", we are in effect considering only the *verifiable* information the manager releases. Big firms will probably enjoy fewer information asymmetries because more analysts follow them rather than due to significant differences in disclosure requirements between, say, the NYSE and the NASDAQ markets.

kind of measure of the speed at which uncertainty is resolved, since it is the correlation coefficient between the return on the risky technology θ and a signal which can be more or less noisy *irrespective of the industry (or firm) considered*. Outsiders, who know the value of ρ_m (but do not observe X_1), are aware that the ρ_b based on which they will make decisions is only the product of ρ_m and ρ_{XY}; hence they rely on it to set a cutoff value Y_1^0, to be enforced once a particular realization y_1 has happened, but accept the fact that the manager has more accurate information and that true values are to be computed using ρ_m (which is common knowledge). The whole Section 3 can be seen as a particular case of this framework when ρ_{XY}=0: outsiders are fed no information whatsoever.

4.2 Designing the Contractual Terms

As of t=0, the entrepreneur can write covenants that will be attached to the securities he sells. As we discussed before, his goal is to maximize the sale price of bonds plus the value of the equity. He may keep all of it, or may sell some or all of it to outside equityholders. In both cases, he has an incentive to maximize the market value of the firm (to minimize the agency costs he will bear). This is equivalent to saying that he will minimize at t=1 the deviation from the optimal investment policy.

The goal of the covenant is then to $\max\limits_{Q\in[0,I]} \frac{1}{r_2} E\left[\left(Q\theta + (I-Q)r_2\right) | Y_1 = y_1\right]$ at t=1. This is done by investing in the risky technology if and only if $y_1 > Y_1^0$ defined, as in the previous section, by: $U(I, \rho_b, Y_1^0) = U(0, \rho_b, Y_1^0)$: at $y_1 = Y_1^0$, the firm as a whole is indifferent between investing in the risky technology or in riskless bonds. However, the decision is made here based on the observation of the signal Y_1, i.e. using the parameters $\rho_b = \rho_{XY}\rho_m$ and $\sigma_{Y_1}^2 = \sigma_{X_1}^2 + \sigma_\varepsilon^2$ instead of ρ_m and $\sigma_{X_1}^2$. Since the cutoff value is increasing in ρ (Theorem 1) and decreasing in the variance of the signal (i.e. σ_{X_1} for X_1^0, σ_{Y_1} for Y_1^0)[36], $Y_1^0 < X_1^0$: the presence of the noise ε makes perfect contracting (and the total elimination of agency costs) not possible.

The problem now is to figure out the terms on which bondholders will contract. We assume here that contracts based on ex-post realizations are ruled out by legal structures. Since $\lim\limits_{\rho_{XY}\to 0} Y_1^0 = \lim\limits_{\rho_m \to 0} X_1^0 = -\infty$, forcing the manager to invest in the risky project if $y_1 > Y_1^0$ may in certain cases (i.e. for low values of ρ_{XY}) add very little value or even be counter-productive, creating new agency costs[37]. At first sight, the following contract could be written by bondholders at t=0:

[36] Proof available upon request from the author.

[37] We remind the reader that in the Myers (1977) framework, enforcing a minimum level of investment may create new agency costs due to forced investment in negative NPV states.

- if $Y_1^0 \geq X_1^F$, impose an investment cutoff of Y_1^0: if $y_1 \leq Y_1^0$, the manager will have to invest in riskless bonds and if $y_1 > Y_1^0$, the risky technology will be chosen.
- if $Y_1^0 < X_1^F$, bondholders will let the shareholder-aligned manager do as she pleases.

However, we can achieve the same result by merely stating that the manager is not allowed to invest below a certain cutoff Y_1^0, whatever the value of the parameters. Moreover, covenants in our framework usually do not specify what the manager has to do above a certain cutoff but merely state what she cannot do[38]. Our two cases hence become:

- if $Y_1^0 \geq X_1^F$, the interdiction of investing in the risky technology below Y_1^0 is binding: the shareholder-aligned manager would want to go down to X_1^F when making her investment decision but is prevented from doing so by the bond covenants;
- if $Y_1^0 < X_1^F$, the interdiction of investing in the risky technology below Y_1^0 is still in effect, but the manager will choose on her own not to invest below X_1^F .

The reason for modifying such a contract is the following: insiders observe both realizations x_1 and ε_1 (the subscript on ε is added to denote a particular realization). Forcing the manager to invest above Y_1^0 (when $Y_1^0 \geq X_1^F$) means that she will invest in the risky technology if and only if $y_1 = x_1 + e_1 > Y_1^0$, i.e. $x_1 > Y_1^0 - e_1$. Now, in the case where $\varepsilon_1 > 0$ and Y_1^0 is only slightly larger than X_1^F, $Y_1^0 - e_1 < X_1^F < Y_1^0$. For $x_1 \in [Y_1^0 - e_1, X_1^F]$, neither shareholders nor bondholders would invest if complete contracting was possible (i.e. if outsiders could observe x_1) and hence *both are worse off if the covenant forces investment whenever* $y_1 > Y_1^0$ *, even when* $Y_1^0 \geq X_1^F$.

This problem is solved if the covenant only states that the manager cannot invest below a certain cutoff Y_1^0 but does not specify what she has to do above the cutoff. The extent of residual risk-shifting after this partial contracting is therefore $X_1^0 - \max(Y_1^0, X_1^F) = \min(X_1^0 - Y_1^0, X_1^0 - X_1^F)$, and in the sequel we will refer to

---

[38] We are grateful to Professor Amihud for pointing this out.

optimal contracting as the bondholders imposing the interdiction to invest whenever their observation y_1 falls short of the cutoff Y_1^0, *no matter how the latter compares to* X_1^F[39].

One may think that new agency costs may be created by contracting under those terms. This is the case if Y_1^0 is very close to X_1^0 and $\varepsilon_1<0$. In that case, bondholders will rule out investing in the risky technology if $y_1 < Y_1^0$, i.e. if $x_1 < Y_1^0 - \varepsilon_1$. However, for $x_1 \in [X_1^0, Y_1^0 - \varepsilon_1]$ (remember that we are considering the case where $\varepsilon_1<0$), one should really invest, but the covenant will not allow it. This, however, is not a problem: Y_1^0 tends to X_1^0 only when ρ_{XY} tends to 1 (we'll shortly show that $\partial Y_1^0 / \partial \rho_{XY} > 0$; it should be obvious that $\lim_{\rho_{XY} \to 1} Y_1^0 = X_1^0$), in which case σ_ε tends to 0 ($\sigma_\varepsilon^2 = \sigma_{X_1}^2 (1-\rho_{XY}^2)/\rho_{XY}^2$). The probability of a realization ε_1 such that $X_1^0 < Y_1^0 - \varepsilon_1$, $\Phi((Y_1^0 - X_1^0)/\sigma_\varepsilon)$, becomes in turn arbitrarily small.

The terms of our covenant, i.e. forbidding investment when the realization y_1 falls short of a certain cutoff Y_1^0, regardless of whether $Y_1^0 > X_1^F$ or $Y_1^0 \leq X_1^F$, thus minimizes the remaining agency costs which are due to the fact that $Y_1^0 < X_1^0$[40].

Proposition 2*: The optimal covenant is the one specifying that investment in the risky technology is forbidden for* $y_1 \leq Y_1^0$.

Proof: this follows from the earlier discussion and the fact that ε is a zero-mean random variable.

Figure 4 shows the two possible cases, i.e. when contracting is interesting and when bondholders cannot reduce the extent of risk-shifting:

[39] It is still of interest to know under what parameters such a contracting is useful (binding), i.e. when $Y_1^0 \geq X_1^F$. Since neither Y_1^0 nor X_1^F admits a closed-form representation, the particular ρ_m^* that equates Y_1^0 and X_1^F for a given σ_ε or ρ_{XY} will have to be solved for numerically.

[40] See also Amihud, Garbade and Kahan (1999) and Smith and Warner (1979).

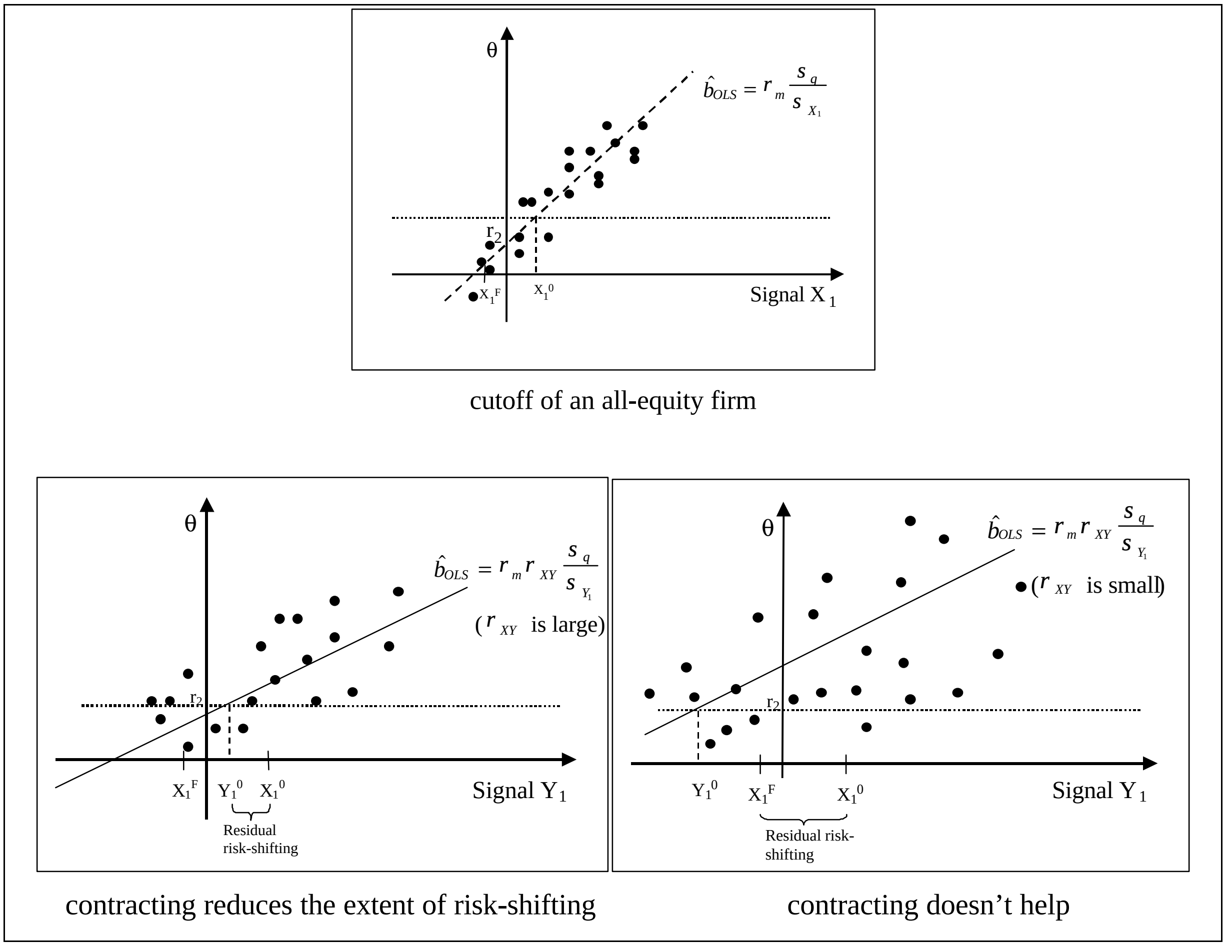


Figure 4

4.3. A Few Preliminary Results

We will now state a few propositions that are more or less obvious, but which are useful both for later results and for intuition's sake.

Proposition 3*: The benefit from contracting is non-decreasing in the face value of the debt outstanding.*

Proof: ◆ if $Y_1^0 \geq X_1^F$ for all F∈[0,Ir$_2$], the average improvement in terms of X$_1$-region due to contracting is $Y_1^0 - X_1^F$. Since $\frac{\partial Y_1^0}{\partial F} = 0$ and $\frac{\partial X_1^F}{\partial F} < 0$, this improvement is strictly increasing in F.

- if $Y_1^0 < X_1^F$, the extent of residual risk-shifting is $X_1^0 - X_1^F$, the same as the original one. However, as F increases, X_1^F decreases and might cross to the left of Y_1^0, yielding a positive improvement.

The intuition for that is pretty straightforward: the precision of the signal is independent of how much debt the firm has issued[41], but risk-shifting becomes more severe the higher the amount of debt outstanding; hence the improvement brought by contracting has to be at least as large for higher debt levels. In particular, there might be no improvement at all for relatively small amounts of debt (Y_1 is too noisy an estimate of X_1), but as the proportion of debt in the capital structure becomes more significant, the improvement becomes positive (Y_1 is still as noisy, but given the extent of the deviation from optimal investment policy, outsiders think twice before discarding the information y_1; in other words, X_1^F crosses to the left of Y_1^0).

Proposition 4*: For a given firm/industry characterized by r_m, improvement in risk-shifting is non-decreasing in r_{XY} (nonincreasing in s_e).*

Proof: $\frac{\partial(Y_1^0 - X_1^F)}{\partial r_{XY}} = \frac{\partial Y_1^0}{\partial r_{XY}} > 0$ (see Appendix 10) or, equivalently, $\frac{\partial(Y_1^0 - X_1^F)}{\partial s_e^2} = \frac{\partial Y_1^0}{\partial s_e^2} < 0$.

Therefore, if $Y_1^0 \geq X_1^F$ for a given F∈[0,Ir$_2$], the improvement increases with ρ_{XY} (decreases with σ_ε); if $Y_1^0 < X_1^F$, the improvement is independent of ρ_{XY} and σ_ε, but as ρ_{XY} increases (as σ_ε decreases), Y_1^0 increases and might cross to the right of X_1^F, yielding a positive improvement. The existence of a unique ρ_{XY} such that $Y_1^0 = X_1^F$ comes from the facts that $\lim_{r_{XY}\to 0} Y_1^0 = -\infty$ and $\lim_{r_{XY}\to 1} Y_1^0 = X_1^0 > X_1^F$ (see Appendix 10), combined with the Intermediate Value Theorem.

The idea behind this is that if the signal observed by outsiders is very noisy, they will be aware of the fact that a good realization of θ might still follow a bad y_1 and prefer to let shareholders conduct a more educated choice than the one they can enforce. As Y_1 becomes more correlated with X_1, outsiders know that they know more and are in a better situation to monitor effectively the investment policy of the firm.

[41] This might not be completely true: there is a secondary effect linked to size: larger firms are likely to issue more debt, and at the same time, as seen previously, to have stricter disclosure requirements (or to be more closely followed by analysts). The precision of the signal is then increasing in the amount of debt issued. But since this is not a direct effect, we decide to ignore it for the time being.

4.4. On the Extent of Residual Risk-Shifting After Optimal Contracting

We can now turn to the question behind our Theorem 2: is it the case that, *after optimal contracting*, the extent of risk-shifting is decreasing in the speed of resolution of uncertainty? In other words, is it the case that a firm whose uncertainty is resolved only late will have more of a tendency to overinvest than a firm whose information comes quickly, once they both have satisfied their disclosure requirements and respected bond covenants? We can indeed answer positively to this question. But before that, we'll show that the benefit at t=1 from contracting, $\max(Y_1^0 - X_1^F, 0)$, is non-increasing in ρ_m:

Theorem 4*: at t=1, the benefit from contracting, i.e. the reduction in the extent of risk-shifiting, is non-increasing in the pattern of resolution of uncertainty ρ_m. More precisely, for ρ_{XY} not too small, i.e. for $\rho_{XY} \geq \underline{\rho_{XY}}$* (or, equivalently, for $\sigma_e \leq \overline{\sigma}_e \equiv \sqrt{\sigma_{X_1}^2 (1 - \underline{\rho_{XY}}^2) / \underline{\rho_{XY}}^2}$), *there exists a unique ρ_m^* such that $Y_1^0 > X_1^F$ for $\rho_m < \rho_m^*$ and $Y_1^0 \leq X_1^F$ for $\rho_m \geq \rho_m^*$. The reduction in the extent of risk-shifting is then decreasing on $[0, \rho_m^*]$ and 0 on $[\rho_m^*, 1]$.*

Proof: see Appendix 10. If $\rho_{XY} < \underline{\rho_{XY}}$, we set $\rho_m^* = 0$: no industry benefits from contracting[42].

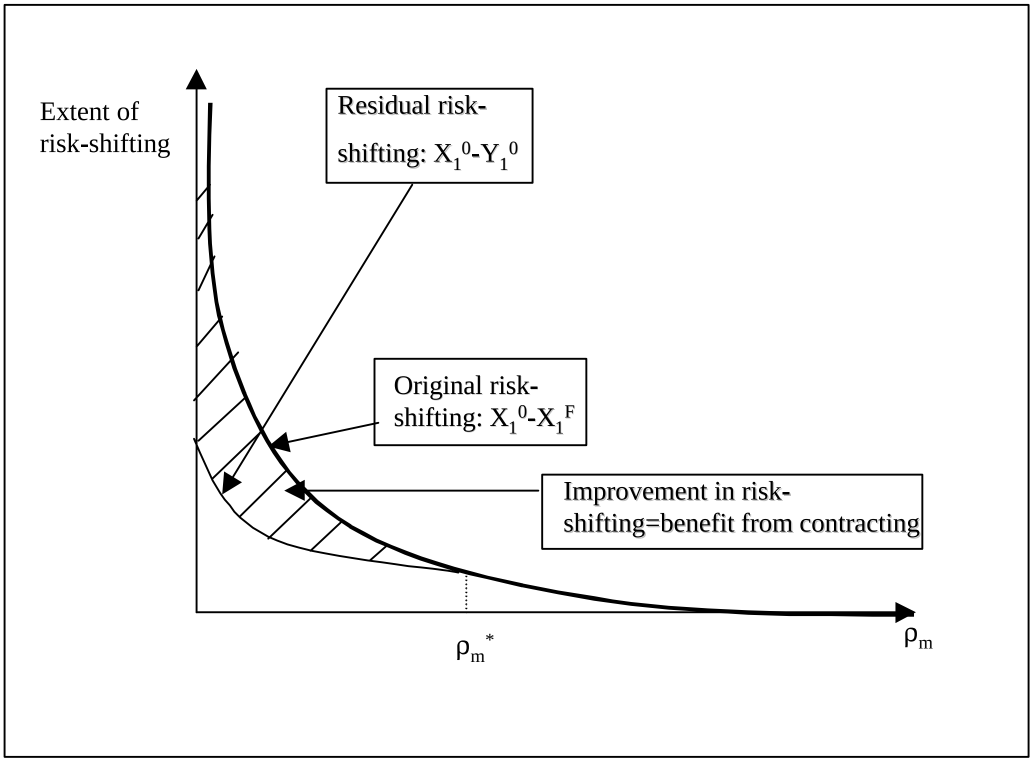
The fact that contracting is more interesting the lower ρ_m is pretty intuitive: a levered firm operating in a field where uncertainty is resolved early does not have, as seen earlier, an investment policy significantly different from its all-equity equivalent. It therefore comes as no surprise that, for a fixed level of disclosure requirement ρ_{XY}, it benefits less from contracting than a firm operating in a field where uncertainty is resolved later and which will have a tendency to risk-shift more significantly. Hence the following corollary:

Corollary 4*: Given a certain strictness of disclosure requirements $\rho_{XY} \geq \underline{\rho_{XY}}$, if a firm/industry benefits from contracting, so will a firm/industry for which uncertainty is resolved only later.*

Proof: this comes directly from Theorem 4 and the fact that for $\rho_{XY} \geq \underline{\rho_{XY}}$, it will be interesting for any firm below ρ_m^* (i.e. for which $Y_1^0 > X_1^F$) to contract. The decreasingness of $Y_1^0 - X_1^F$ in ρ_m over the region $[0, \rho_m^*]$ ensures that on this region, the lower ρ_m, the higher the reduction in the extent of risk-shifting.

[42] As we noted earlier, a lower ρ_{XY} (signal Y_1 is less correlated to signal X_1) is equivalent to a higher σ_ε (signal Y is a more noisy version of X_1). In the sequel, we'll mainly do comparative statics with respect to ρ_{XY}, since we have a more intuitive feeling for it (strictness of disclosure requirements), and also because σ_ε has to be compared to σ_{X_1} to have any sense at all.

The following figure illustrates the case where $r_{XY} \geq \underline{r_{XY}}$:



<u>Figure 5</u>

Corollary 4 yields a r_{XY} / r_m^* "frontier" that will give us the r_m^* cutoff (below which firms benefit from contracting, above which contracting is useless) as a function of ρ_{XY}. Since r_m^* is shown in Appendix 10 to be increasing in ρ_{XY}, we get the following corollary as a bonus:

> ***Corollary 4 bis****: The economy as a whole is better off, i.e. more industries benefit from contracting, the higher* r_{XY}.

This is not as trivial as it may seem: it is pretty intuitive that the stricter the disclosure requirements, the more each firm/industry benefits from contracting. However, we state that on top of that, an increase in ρ_{XY} will bring new industries to contract; these are the industries that were just a little bit too mature (ρ_m too high) to benefit from a disclosure of quality ρ_{XY}, but who can gain from writing contracts based on information of quality $\rho_{XY}+\Delta\rho_{XY}$. In the limit, when ρ_{XY} tends to 1 (outsiders know as much as insiders), the optimal cutoff

X_1^0 can be imposed by bondholders, leaving zero agency costs: *regardless of the amount of debt outstanding and the industry in which it operates, a firm will benefit from contracting*. At the opposite end of the spectrum, when ρ_{XY} crosses $\underline{r}_{XY}$ to the left, *no firm will benefit from contracting, no matter how large the amount of debt outstanding and regardless of the industry in which it operates*.

However, if we keep ρ_{XY} constant, we are still left with the equivalent of Theorem 2, once optimal contracting has been written:

Theorem 2 bis*: The extent of residual risk-shifting* $X_1^0 - \max(Y_1^0, X_1^F)$ *is decreasing in the speed of resolution of uncertainty* r_m.

Proof: i) for firms with $r_m < r_m^*$, we want to show that $\partial(X_1^0 - Y_1^0)/\partial r_m < 0$; since $X_1^0 = \lim_{r_{XY}\to 1} Y_1^0$, one way of doing it is to show that $\partial^2 Y_1^0 / \partial r_{XY} \partial r_m < 0$, thereby ensuring that $\frac{\partial Y_1^0}{\partial r_m} > \frac{\partial X_1^F}{\partial r_m}$. This proof is cumbersome and is not reported in the Appendix, but is available upon request from the author;

ii) for firms with $r_m \geq r_m^*$, $\partial(X_1^0 - X_1^F)/\partial r_m < 0$ as seen in Theorem 2[43].

Partial contracting will thus mitigate risk-shifting, but will not destroy the decreasingness of the extent of risk-shifting in the pattern of temporal resolution of uncertainty. Thus all our discussion following Theorem 2 is still valid after bondholders contract as well as they can.

The reduction in agency costs is not monotonic in ρ_m. This is due, as we saw before, to the fact that for low values of ρ_m, the manager might deviate a lot at t=1 from the optimal investment policy, but the quality of his information is so poor that this will not have serious consequences in terms of agency costs as of t=0. Agency costs are then virtually equal to zero and cannot be reduced by contracting. At the opposite end of the ρ_m spectrum, when ρ_m tends to 1, the manager's policy is essentially the same as the value-maximizing one. However, we argue that our comparative statics on the residual extent of risk-shifting at t=1, $\min(X_1^0 - Y_1^0, X_1^0 - X_1^F)$, are still of interest: the goal of a contract is to make sure that the investment policy of the firm will not deviate consequently from the optimal one, and any reduction in this deviation at t=1 will reduce agency costs, albeit not monotonically, at t=0.

---

[43] For $r_m = r_m^*$, the fact that $\partial(Y_1^0 - X_1^F)/\partial r_m$ is still negative when $Y_1^0 = X_1^F$ (see Appendix 10) implies that the residual risk-shifting, $\min(X_1^0 - Y_1^0, X_1^0 - X_1^F)$, is not differentiable in ρ_m at $r_m = r_m^*$; however, we are not worried about non-differentiability on a set of measure zero.

4.5. On Bond Yields After Contracting

Our last task is to investigate how bond yields evolve with the temporal resolution of uncertainty after bondholders impose the optimal cutoff $\max(X_1^F, Y_1^0)$. It should be clear from the discussion in Appendix 11 that the way we compute bond prices remains the same[44], with the difference that the cutoff Y_1^0 is substituted for X_1^F whenever it is greater. We show there the following:

Theorem 3 bis: *Once bondholders have contracted as well as they can, the equilibrium prices of corporate bonds are still increasing in r_m, the pattern of temporal resolution of uncertainty. Equivalently, the default premium demanded on corporate bonds is still decreasing in r_m.*

As before, this stems from two effects: i) the "total firm effect" remains, i.e. the manager can carry out a more "educated" investment policy (alternatively, the residual variance of the risky project, if it is entered into, is lower: $\text{var}(q \mid X_1 = x_1) = s_q^2(1 - r_m^2)$), the higher ρ$_m$; bondholders share this benefit with shareholders[45] and ii) as ρ$_m$ increases, the extent of risk-shifting $X_1^0 - \max(X_1^F, Y_1^0)$ decreases (see Theorem 2 bis), and bondholders benefit from a lower deviation from the socially optimal investment policy.

Finally, bond yields evolve monotonically in the strictness of disclosure requirements:

Theorem 5: *Bond prices are non-decreasing in the strictness of disclosure requirements r_{XY} (non-increasing in the report noise s_e). Equivalently, the default premium demanded on corporate bonds is non-increasing in r_{XY} (non-decreasing in the report noise[46]). More precisely, for r_{XY} not too small, i.e. for $r_{XY} \geq \underline{r_{XY}}$, bond yields are decreasing in r_{XY} on $[0, r_m^*]$ and independent of r_{XY} on $[r_m^*, 1]$.*

Proof: see Appendix 11.

---

[44] As we noted earlier, bondholders will use ρ$_m$, which is public knowledge, to price all assets.

[45] As noted in Section 4 and proved in Appendix 8, firm value is increasing in ρ$_m$ as long as the cutoff the firm uses for investment is no larger than X_1^0. Since this is the case for both X_1^F and Y_1^0, it is also the case for $\max(X_1^F, Y_1^0)$ and firm value is still increasing in ρ$_m$.

[46] or, equivalently, non-decreasing in the amount of asymmetric information.

Here, there is no "total firm value effect": the decrease in bond yields accompanying an increase in ρ_{XY} is solely due to the fact that the extent of residual risk shifting $X_1^0 - \max(X_1^F, Y_1^0)$ is decreasing in ρ_{XY} whenever $Y_1^0 > X_1^F$, i.e. when $r_{XY} \geq \underline{r_{XY}}$ and $r_m < r_m^*$ (following Proposition 4) and bondholders, rationally anticipating a lower deviation from optimal investment policy, demand a lower yield premium.

This feature of our model is consistent with the existing literature on the effect of accounting reports on security prices. Most recently, Duffie and Lando (1998) refine the Leland (1994) model of default and allow for "imperfect information", i.e. outsiders observe only at discrete times a noisy accounting report on the value of assets, and survivorship. Their conclusion is that the zero-coupon credit spread is strictly increasing in the reporting noise level *a* (see their Figure 7). In our model, reports could be not only about accounting data, but also about the investment opportunity set facing the firm (how encouraging a signal observed at t=1 is). Since the quantity that characterizes the noisiness of this report in our model is σ_ε, our conclusion is very similar to Duffie and Lando's one.

It is worth noting that share prices are decreasing in ρ_{XY} (see Appendix 11): although an increase in ρ_{XY} will increase firm value, this benefits bondholders and hurts shareholders, prevented from investing on the region $[Y_1^0, X_1^0]$. Equityholders rationally anticipate that their manager won't have the freedom to maximize the price of their claim and will drive the latter down. Note finally that we constrained ρ_{XY} to be the same across industries (hence our intuition that varying ρ_{XY} will yield *intra-industry* comparative statics). It would be of interest to let ρ_{XY} vary across industries. For instance, by positing that $r_{XY} = r_m^{n-1}$ or, equivalently, $r_b = r_m^n$ for n≥2, we are in effect imposing the restriction that there is more information asymmetry in industries where uncertainty is resolved later, which seems to be the intuitive case. We looked at the matter but did not report our results, which remain essentially unchanged, here.

5. Empirical Implications and Empirical Evidence

The different lemmas and theorems derived in this paper have empirical implications, some of which are more readily testable than others. The main problem will be of course to find a good proxy for the speed at which uncertainty is resolved. It could be the time-series correlation between forecasted earnings and realized earnings (or between innovations in earnings E_t-E_{t-1} and forecasted innovations $F_{t-1,t}$), the average forecasting error (or more precisely a root mean square error) or the dispersion in the earnings forecasts across analysts. Of course, all these proxies should estimate how much more difficult it is to forecast the long-term than it is to forecast the short term. We refer the reader to our companion paper (Reisz (1999)), which uses those proxies. It seems that related proxies (in the existing literature) are the amount of R&D and advertising expenses (scaled by sales or market value of assets) as well as how job-specific the skills of the workers employed by the firm are[47]. The empirical implications are presented in the order in which they appear in our paper.

a) From Theorem 1, b), we would expect firms operating in fields where uncertainty is resolved later to have more risky cash flows. This effect should be translated as well in the variance of earnings as in the variance of total assets in place (although firms that rely heavily on R&D and goodwill will have a lot of assets that do not appear in the books) or even in the unlevered beta[48]. We also expect these firms to invest more (for instance through setting a lower hurdle rate for their investments) on a size-corrected basis. It is worth noting that Bradley, Jarrel and Kim (1984) as well as Titman and Wessels (1988) found a non-significant *negative* correlation between R&D expenses and the volatility of the firm, which contradicts not only our Theorem 1, but our basic intuition as well. This may be due to the proxy for the volatility of a firm's operations used in those studies. Bradley, Jarrell and Kim (1984) look at the standard deviation of the first differences in annual earnings before depreciation, interest and taxes, while Titman and Wessels (1988) look at the standard deviation of the percentage change in operating income. Both proxies ignore the effect of taxes on cash flows (it could be argued that firms with more delayed resolution of uncertainty pay fewer taxes since a larger fraction of their value is accounted by growth options), as well as the widely reported phenomenon of “earnings management”.

b) Since risk shifting as of t=1 is a more severe problem for a lower ρ (i.e. the later the temporal resolution of uncertainty) according to both our Theorem 2 and Theorem 2 bis, we would expect to observe the following empirical evidence: firms operating in a field where uncertainty is resolved quite late in the life of the project (e.g. R&D intensive firms) should be reluctant to issuing risky debt, since investors will rationally anticipate severe risk shifting, as opposed to firms operating in fields where uncertainty is resolved very quickly (e.g. a timber firm whose investment in a forest may experience great uncertainty in the early years of the trees’ growth, but once this uncertainty is resolved, the future cash flows may be more or less predictable) and where risk shifting will not be so severe. This effect is reinforced by the fact that firms facing late temporal resolution of uncertainty have more risky operations (Theorem 1), and this added variance decreases the debt capacity of the firm.
   - This is consistent with Bradley, Jarrell and Kim (1984) who report a “strong finding of intra-industry similarities in firm leverage ratios and of persistent inter-industry differences” in a way that supports our model: “54% of the cross-sectional variance in firm leverage ratios can be explained by industrial

---

[47] With the problem that those proxies control also for other effects, such as the importance of growth options in the investment opportunity set of a particular firm.

[48] The unlevered beta proxies for the *systematic* risk of the firm; why a firm operating in a field where uncertainty is resolved only late should have a higher *systematic* (and not only overall) risk comes from the fact that we assumed in our model that all firms have the *same* risky technology available (see it as a market risk) and that the ones with low ρ will tend to invest more often.

classification, with more variation in mean leverage ratios across industries than there is in firm leverage ratio within industries". Industries like "Drugs and Cosmetics", "Electronics" or "Petroleum Exploration", which we would heuristically characterize as having later temporal resolution of uncertainty, have significantly lower debt levels than industries like "Steel", "Telephone", "Electricity and Gas Utilities" or "Airlines", which seem to have earlier temporal resolution of uncertainty.[49]. There is an argument that could be made against these results: the firms that we characterize as having later temporal resolution of uncertainty are also firms that seem to have higher overall risk (and that do, according to our first theorem) and therefore their lower debt levels could only be due to this higher risk. However, Bradley, Jarrell and Kim (1984) regress debt to value ratio on firm volatility, non-debt tax shields and R&D and advertising expenses (their Table III), *thereby controlling for the overall volatility of assets as well as for the non debt tax shield due to R&D expenses*. The coefficient on R&D and advertising expenses is significantly negative, as our model predicted (firms that have high R&D expenses are usually firms in a late temporal resolution of uncertainty field and should therefore be more reluctant to issuing debt).

- A more comprehensive test is carried out by Long and Malitz (1986) who look at 545 manufacturing firms, *grouped into portfolios in order to hold the operating or business risk (as measured by unlevered betas) of all firms constant*. The quartile with highest R&D expenses is the one with lowest leverage and the one with lowest R&D expenses is the one with highest leverage. The negative relationship between R&D (and advertising) expenses and leverage remains when the latter is regressed on the former, controlling for the firm systematic and residual risk and non-debt tax shields[50]. This is a summary of previous work of theirs where they looked at 63 industries classified by a four-digit SIC code; of the five industries with the lowest leverage, four have the highest R&D and advertising expenditures, while the five industries with the highest leverage show the lowest percentage of these intangible investments.
- Similar results are reported by Titman and Wessels (1988) who use a factor-analytic technique, linear structural modeling, that mitigates the measurement problem encountered when working with proxy variables. They report that firms that are more "unique" tend to have lower debt/equity ratios. They define "uniqueness" as higher R&D and advertising expenses and lower quit ratios ("firms that sell products with close substitutes are likely to do less research and development since their innovations can be more easily duplicated" whereas "firms with relatively unique products are expected to advertise more and, in general, spend more in promoting and selling their products" and "employ workers with high levels of job-specific human capital who will thus find it costly to leave their jobs"[51]). Intuition dictates us that firms that operate in an industry where uncertainty is resolved late

---

[49] See their Table I p. 870. Their results are robust to the exclusion of regulated industries.

[50] Their results carry over as well for portfolios as for individual firms.

[51] p. 5. They also control for the tax deductibility of R&D and some selling expenses. Titman (1984) also finds that "firms that can potentially impose high costs on their customers, workers and suppliers in the event of liquidation have lower debt

are, in a Titman and Wessels sense, more "unique"[52] and in that heuristic sense, our model is supported by their empirical evidence.

- Finally, this discrepancy of capital structure across industries is in no way particular to the United States. Kester (1986), for example, reports that industry dummies are highly significant in the case of Japanese capital structure, especially for "debt-equity ratios in the upper quartile of the sample. Most of these are mature, heavy industries and include steel, general chemicals, nonferrous metals, paper and petroleum refining" (p. 12), all of which correspond to fields where uncertainty is resolved relatively early. Bronte (1982) reports that after netting out debt refinancing, internal sources accounted for 50.5% of net capital invested in Japan in 1970 and 102.4% in 1979, indicating that internally generated cash was being used to retire debt over that period, with the trend "most pronounced in such high technology industries as electronics, pharmaceuticals and communication equipment"[53].

It is worth mentioning that the negative relation between leverage and TRU is still present when more direct proxies for TRU are used (see Reisz (1999)).

c) If firms operating in a low ρ environment have already decided to take on debt (for instance because insiders could not bring enough capital on their own or because the opportunity cost of inside funds or outside equity are even higher), we should see more representatives of creditors (bondholders or lending banks) sitting on the board of directors -- so as to prevent shareholders from indulging too much in their risk shifting incentives -- than for firms operating in a higher ρ environment. Now, since the latter will typically have more risky debt outstanding, tests should control for it and investigate whether low ρ firms have more creditors, *per dollar lent*, on the board of directors.

d) Bondholders, supposed to rationally anticipate risk shifting incentives, will demand a higher yield from a corporation in a low ρ industry than from a corporation in a high ρ field (Theorems 3 and 3 bis). It should be investigated whether it is the case once risk is already accounted for, since as we showed in Theorem 1, firms for which uncertainty is resolved later are also riskier. Our companion paper (Reisz (1999)) answers positively to the question of whether temporal resolution of uncertainty still has some power in explaining cross-sectional variation in bonds yields once risk is controlled for.

---

ratios". This definition corresponds, once more heuristically, to our perception of a late temporal resolution of uncertainty pattern.

[52] As, for instance, computer or pharmaceutical firms, whereas firms that operate in a field where uncertainty is resolved earlier tend to offer products that are more substitutable and less characteristic of a given firm and to rely less heavily on human capital.

[53] *Ibid*, p. 15. See also Exhibit 2, p. 9.

e) Since risk-shifting is more of a problem for firms facing delayed resolution of uncertainty, a solution may be to shorten debt maturity for those firms, rendering suboptimal investment by a shareholder-aligned manager impossible. This is also in the best interest of the firms with most delayed resolution of uncertainty, which may be able to refinance on better terms at the intermediate date (in the limit, when no investment is made, lenders will charge the riskless rate). We therefore expect firms with more delayed resolution of uncertainty to issue shorter debt. This is documented by Stohs and Mauer (1996) and Guedes and Opler (1996) if we are willing to see R&D sclaed by market value of assets as a proxy for TRU, and by Reisz (1999) who uses more direct proxies.

f) All of the above empirical implications should be mitigated when the existing regulation enforces verifiable disclosure requirements. In particular, this means that larger firms (or, more generally, firms listed on a stock exchange enforcing stricter disclosure regulations) should display a lower extent of risk-shifting than smaller firms and bondholders should demand lower yield premia from them[54]. This gives us therefore intra-industry empirical implications. It is also worth noting that it might as well yield inter-country comparative statics: since the accounting and reporting standards are typically stricter in the US than in other countries, one might believe that the aforementioned empirical implications would be more obviously displayed in other countries[55]. Tests designed at checking whether the aforementioned empirical implications hold should therefore control for firm size, the stock exchange on which a particular stock is quoted or the country where a particular corporation is registered.

The main difficulty in all future tests to be conducted will be to single out the effect of earlier or later temporal resolution of uncertainty from other effects: firms that have high R&D expenses or low quit ratios, i.e. firms we heuristically perceived as operating in a field where uncertainty is resolved later, are also firms

- that have more intangible assets and therefore higher expected bankruptcy costs (if they go bankrupt, they have more to lose, including the opportunity to further invest in R&D);
- that are more risky, and therefore have not only higher bankruptcy costs in case of default, but also a higher probability of default and, following, higher risk-shifting incentives;
- the value of which is mainly accounted for by expected future earnings (their current earnings might even be negative!), and therefore finding themselves in lower tax brackets.

All these effects contribute in predicting lower debt ratios and higher bond yields for firms operating in fields where uncertainty is resolved only late (i.e. firms having high R&D expenses) *based only on existing theories of bankruptcy costs and tax considerations* as determinants of capital structures[56]. Further empirical work will

[54] With the problem that size may proxy for the liquidity of the firm's bonds as well.

[55] We thank David Yermack for pointing out this inter-country implication.

[56] For instance, our conclusion that firms in a lower ρ field should rely less heavily on debt could be merely due to tax effects: progressivity in the tax structure implies that greater volatility in taxable income raises the firm’s expected tax

therefore have to tackle the non-trivial problem of isolating the effect of temporal resolution of uncertainty on capital structure, *once the effect of classical determinants has been singled out*. This will be done by choosing a direct proxy for TRU.

However, this does not undermine the validity of our theoretical model and its empirical implications: we come to the same conclusions as existing theories of bankruptcy costs and tax considerations, but *without even considering these effects, therefore in effect singling out the effect of temporal resolution of uncertainty on investment policies and bond yields*: we offer a new element of explanation to the existing empirical evidence, and once we add the aforementioned considerations, our results will only be reinforced. Temporal resolution of uncertainty is thus to be added to tax-related benefits of debt and expected bankruptcy costs in any further analysis of capital structure and corporate debt yields to acquire a fuller understanding of the available empirical evidence.

6. Discussion and Concluding Remarks

In the tradition of papers dealing with the interactions between product markets and financial decisions, we argued that a firm's financial decisions *cannot be independent* of how quickly uncertainty is resolved in the field in which the firm operates. It was already a well-known fact that a firm which has risky debt outstanding will suffer from the risk-shifting incentives of a shareholder-aligned manager (with the extent to which the levered firm will overinvest being increasing in the amount of risky debt outstanding). Our paper's innovation is to show that firms for which uncertainty will be resolved later will suffer far more from those overinvestment tendencies. As a result, the yield premium demanded on corporate bonds will be higher the later the uncertainty is resolved. This is only mitigated by partial contracting, but the qualitative conclusions remain the same.

These results add a new insight to the conclusions of classical bankruptcy costs / tax clientele theories and should therefore only be reinforced by the inclusion of those considerations in a more general model. Nonetheless, we see the inclusion of offsetting agency costs of equity as a more interesting task to undertake in order to determine an optimal capital structure as a function of temporal resolution of uncertainty.

The fact that firm value as well as bond prices are higher the earlier the uncertainty is resolved does not, however, invalidate Ross' (1989) theory of resolution irrelevancy since the difference in prices in our model

---

liabilities (see Smith and Stulz (1985)). Our Theorem 1, stating that firms with lower ρ have a higher variance of final cash flows, implies that theses firms have an incentive to reduce the amount of debt in their capital structure over the range of progressivity.

are due to the anticipation of different managerial behaviors that can lead to different cash flows. In Ross' words, changing the pattern of temporal resolution of uncertainty alters the no-arbitrage martingale pricing operator (which is equivalent to changing the state space spanning). As a result, *we do not even need risk-aversion for the speed of resolution of uncertainty to affect security prices*.

Further work should investigate the optimal financing of different projects based on how quickly uncertainty is resolved. This encompasses as well the nature of the claims to be used as their maturity; a richer menu of contractual forms for outside claims (for instance convertible debt) should also be introduced. It should also try to explain the cross-sectional differences in dividend policies and add a "temporal" explanation to why mature firms differ markedly in this respect from growth firms. Another avenue for further research is to investigate how the compensation package offered to the top officers of levered corporations should depend on the pattern of resolution of uncertainty in the field in which the firm operates[57], as well as to design binding commitments with suppliers and contractors that would impose a financial penalty upon the firm if the better project is not undertaken. This should not only document the wide differences in managerial compensation across industries, but also suggest a way of driving managerial decisions closer to the optimal investment choices as a function of the industry in which the firm operates.

[57] Keeping in mind that the manager might not be shareholder-aligned in the first place. This may be due, for instance, to the fact that the executive's time horizon is relatively short while the value of the stock is the present value of dividends stretching to infinity. See Chidambaran and John (1998).

Appendix

1. A Reminder on Multivariate Normal Distributions

Let Y and X be vector random variables normally distributed with mean vector [EX EY]' and variance-covariance matrix Ω. We will partition Ω as:

$$\Omega = \begin{bmatrix} \Omega_{XX} & \Omega_{XY} \\ \Omega_{YX} & \Omega_{YY} \end{bmatrix}$$

Then Y conditional on X=x is also normally distributed with mean $EX+\Omega_{YX}\Omega_{XX}^{-1}(x-EX)$ and variance-covariance matrix $\Omega_{YY}-\Omega_{YX}\Omega_{XX}^{-1}\Omega_{XY}$. In particular, if two scalar random variables X and θ are jointly normally distributed, then the distribution of θ conditional on X is also normal with the following parameters:

$$E(\theta \mid X=x) = E(\theta) + \frac{\operatorname{cov}(X,\theta)}{\operatorname{var}(X)}\big(x - E(X)\big) = E(\theta) + \rho\,\frac{\sigma_\theta}{\sigma_X}(x - E(X))$$

$$\operatorname{var}(\theta \mid X) = \operatorname{var}(\theta).(1-\rho^2)$$

where ρ stands for the correlation between X and θ and σ_θ and σ_X are the respective standard deviations of θ and X. For more details, see Feller (1976) or Goldberger (1964).

2. Computing the Objective Function

From here on, to simplify notations, we will write $E(\theta|X_1)$ for $E(\theta|X_1=x_1)$ and $\operatorname{var}(\theta|X_1)$ for $\operatorname{var}(\theta|X_1=x_1)$; however, our expressions, as for instance $U(I,\rho,X_1,F)$, depend on a *particular realization x_1*. Equation (1) can be expressed as:

$$\int_{\theta^*}^{\infty} \big[Q\theta + (I-Q)r_2 - F\big] \frac{1}{\sqrt{2\pi \operatorname{var}(\theta \mid X_1)}} \exp\left\{-\frac{[\theta - E(\theta \mid X_1)]^2}{2\operatorname{var}(\theta \mid X_1)}\right\} d\theta$$

$$= \big[(I-Q)r_2 - F\big]\Phi\left\{\frac{E(\theta \mid X_1) - \theta^*}{\sqrt{\operatorname{var}(\theta \mid X_1)}}\right\} + Q\int_{\theta^*}^{\infty} \frac{\theta}{\sqrt{2\pi \operatorname{var}(\theta \mid X_1)}} \exp\left\{-\frac{[\theta - E(\theta \mid X_1)]^2}{2\operatorname{var}(\theta \mid X_1)}\right\} d\theta$$

Now, the second term is equal to, after a change of variable making *z* a standard normal variate:

$$Q\int_{\frac{q^*-E(q|X_1)}{\sqrt{\operatorname{var}(q|X_1)}}}^{\infty}\left[E(q\,|\,X_1)+z\sqrt{\operatorname{var}(q\,|\,X_1)}\right]\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\,dz$$

$$=QE(q\,|\,X_1)\Phi\left(\frac{E(q\,|\,X_1)-q^*}{\sqrt{\operatorname{var}(q\,|\,X_1)}}\right)+Q\sqrt{\operatorname{var}(q\,|\,X_1)}\left[\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\right]_{z=\frac{q^*-E(q|X_1)}{\sqrt{\operatorname{var}(q|X_1)}}}^{z=\infty}$$

$$=QE(q\,|\,X_1)\Phi\left(\frac{E(q\,|\,X_1)-q^*}{\sqrt{\operatorname{var}(q\,|\,X_1)}}\right)+Q\sqrt{\operatorname{var}(q\,|\,X_1)}\,\varphi\left(\frac{q^*-E(q\,|\,X_1)}{\sqrt{\operatorname{var}(q\,|\,X_1)}}\right)$$

A final arrangement yields the expression in the text.

3. Computing the First and Second Derivatives

From equation (2),

$$U(Q,r,X_1,F)=\left[Ir_2+Q\left[E(q\,|\,X_1)-r_2\right]-F\right]\Phi(B)+Q\sqrt{\operatorname{var}(q\,|\,X_1)}\,\varphi(B)\quad\text{where}\quad B=\frac{E(q\,|\,X_1)-\frac{F-(I-Q)r_2}{Q}}{\sqrt{\operatorname{var}(q\,|\,X_1)}}.$$

Therefore

$$\frac{\partial U(Q,r,X_1,F)}{\partial Q}=\left[E(q\,|\,X_1)-r_2\right]\Phi(B)+\left[Ir_2+Q\left[E(q\,|\,X_1)-r_2\right]-F\right]\varphi(B)\frac{\partial B}{\partial Q}+\sqrt{\operatorname{var}(q\,|\,X_1)}\,\varphi(B)+Q\sqrt{\operatorname{var}(q\,|\,X_1)}\,\frac{\partial\varphi(B)}{\partial Q}$$

now, $\dfrac{\partial\varphi(B)}{\partial Q}=-B\varphi(B)\dfrac{\partial B}{\partial Q}$ and $\dfrac{\partial B}{\partial Q}=\dfrac{F-Ir_2}{Q^2\sqrt{\operatorname{var}(q\,|\,X_1)}}$.

Hence,

$$\frac{\partial U(Q,r,X_1,F)}{\partial Q}=\left[E(q\,|\,X_1)-r_2\right]\Phi(B)+\left[Ir_2+Q\left[E(q\,|\,X_1)-r_2\right]-F\right]\varphi(B)\frac{F-Ir_2}{Q^2\sqrt{\operatorname{var}(q\,|\,X_1)}}$$

$$+\sqrt{\operatorname{var}(q\,|\,X_1)}\,\varphi(B)-Q\sqrt{\operatorname{var}(q\,|\,X_1)}\,B\varphi(B)\frac{F-Ir_2}{Q^2\sqrt{\operatorname{var}(q\,|\,X_1)}}$$

$$=\left[E(q\,|\,X_1)-r_2\right]\Phi(B)+\sqrt{\operatorname{var}(q\,|\,X_1)}\,\varphi(B)$$

Differentiating this once more yields:

$$\frac{\partial^2U(Q,r,X_1,F)}{\partial Q^2}=\left[E(q|X_1)-r_2\right]\varphi(B)\frac{F-Ir_2}{Q^2\sqrt{\operatorname{var}(q\,|\,X_1)}}-\sqrt{\operatorname{var}(q\,|\,X_1)}\,B\varphi(B)\frac{F-Ir_2}{Q^2\sqrt{\operatorname{var}(q\,|\,X_1)}}$$

$$=\varphi(B)\frac{F-Ir_2}{Q^2\sqrt{\operatorname{var}(q\,|\,X_1)}}\left[E(q\,|\,X_1)-r_2-\left[E(q\,|\,X_1)-r_2\right]-\frac{Ir_2-F}{Q}\right]$$

$$=\varphi(B)\frac{\left(F-Ir_2\right)^2}{Q^3\sqrt{\operatorname{var}(q\,|\,X_1)}}>0$$

4. Proof of Lemma 1

As we saw in section 3.2, the manager of a levered firm will essentially compare $U(I,\rho,X_1,F)$ with $U(0,\rho,X_1,F)$ in deciding whether to invest or not. We therefore consider here the quantity $\Delta U(\rho,X_1,F) \equiv U(I,\rho,X_1,F) - U(0,\rho,X_1,F)$. Now $\lim_{Q\to 0} B = +\infty$; hence $\lim_{Q\to 0} U(Q,\rho,X_1,F) = Ir_2 - F$ (note the continuity of U when Q tends to 0: as we noted earlier, if Q=0, the shareholders are left with Ir_2-F after investing everything in riskless Treasury bills and repaying bondholders) and

$$\Delta U(\rho, X_1, F) = \left[IE(\theta|X_1) - F\right]\Phi\left(\frac{IE(\theta|X_1)-F}{I\sqrt{\operatorname{var}(\theta|X_1)}}\right) + I\sqrt{\operatorname{var}(\theta|X_1)}\,\varphi\left(\frac{IE(\theta|X_1)-F}{I\sqrt{\operatorname{var}(\theta|X_1)}}\right) - Ir_2 + F$$

$$\equiv I\sqrt{\operatorname{var}(\theta|X_1)}\left[C\Phi(C) + \varphi(C)\right] - Ir_2 + F \quad \text{where} \quad C \equiv \frac{IE(\theta|X_1)-F}{I\sqrt{\operatorname{var}(\theta|X_1)}} \qquad \text{(A1)}$$

<u>NB</u>: It is worth saying a few words on this expression C (which is the expression B when Q=I): conditional on investing in the risky technology, $\Phi(C)$ is nothing else than the t=1 probability of being solvent at t=2: after investing in the risky technology, bankruptcy occurs if the final cash flow $I\theta$ is less then the promised payment F, which happens with the probability $\int_{-\infty}^{F/I} dP(\theta|X_1)$; this in turn is equal to $\Phi((F/I - E(\theta|X_1))/\sigma_{\theta|X_1}) = 1 - \Phi(C)$ given that θ, conditional on X_1, follows a normal law with mean $E(\theta|X_1)$ and variance $\sigma^2_{\theta|X_1}$.

Differentiating $\Delta U(\rho,X_1,F)$ with respect to X_1 yields:

$$\frac{\partial \Delta U(\rho, X_1)}{\partial X_1} = I\sqrt{\operatorname{var}(\theta|X_1)}\left[\Phi(C)\frac{\partial C}{\partial X_1} + C\varphi(C)\frac{\partial C}{\partial X_1} - C\varphi(C)\frac{\partial C}{\partial X_1}\right]$$

$$= I\rho\sqrt{\frac{\operatorname{var}(\theta)}{\operatorname{var}(X_1)}}\,\Phi(C) > 0 \quad \text{since} \quad \frac{\partial C}{\partial X_1} = \frac{\rho}{\sqrt{(1-\rho^2)\operatorname{var}(X_1)}} \qquad \text{(A2)}$$

Therefore $\Delta U(\rho,X_1)$ is strictly increasing in X_1. Since $\lim_{X_1\to-\infty} \Delta U(\rho, X_1) = -Ir_2 + F < 0$ and $\lim_{X_1\to+\infty} \Delta U(\rho, X_1) = +\infty$ and given the continuity of $\Delta U(\rho,X_1)$ in X_1, there is a unique X_1^F satisfying $\Delta U(\rho,X_1^F)=0$ by the Intermediate Value Theorem. QED.

5. Proof of Lemma 2

To prove that $X_1^0 > X_1^F$ for any positive amount of risky debt F_2, it suffices to show that X_1^F is decreasing in F. Hence any positive value of F will yield a cutoff X_1^F strictly smaller than X_1^0. Now, X_1^F is defined as the value of X_1 for which $\Delta U(X_1,\rho,F)=0$. Implicit differentiation yields, keeping $\Delta U(X_1^F,\rho,F)$ equal to 0:

$$\frac{\partial X_1^F}{\partial F} = \left[\frac{-\dfrac{\partial \Delta U(\rho, X_1, F)}{\partial F_2}}{\dfrac{\partial \Delta U(\rho, X_1, F)}{\partial X_1}} \right]_{X_1 = X_1^F}$$

Now,

$$\begin{aligned}\frac{\partial \Delta U(\rho, X_1, F)}{\partial F} &= I\sqrt{\mathrm{var}(\theta \mid X_1)}\left[\frac{\partial C}{\partial F}\Phi(C) + C\varphi(C)\frac{\partial C}{\partial F} - C\varphi(C)\frac{\partial C}{\partial F}\right] + 1 \\ &= I\sqrt{\mathrm{var}(\theta \mid X_1)}\,\Phi(C)\frac{-1}{I\sqrt{\mathrm{var}(\theta \mid X_1)}} + 1 = 1 - \Phi(C) > 0 \qquad \text{(A3)}\end{aligned}$$

hence, using equation (A2) and (A3), we get:

$$\frac{\partial X_1^F}{\partial F} = \frac{\Phi(C^{X_1^F}) - 1}{I\rho\sqrt{\dfrac{\mathrm{var}(\theta)}{\mathrm{var}(X_1)}}\,\Phi(C^{X_1^F})} < 0 \qquad \text{(A4)}$$

where $C^{X_1^F}$ denotes the expression C taken in $x_1=X_1^F$. QED.

6. Proof of Theorem 1

We will prove here that the earlier the temporal resolution of uncertainty (i.e. the higher ρ), the higher the cutoff value X_1^0 (resp. X_1^F) the manager of an all-equity (resp. levered) firm uses. In the sequel, in order to make equations more readable, we will adopt the following notations: $E(X_1) = \overline{X}_1$, $E(\theta) = \overline{\theta}$, $\sqrt{\mathrm{var}(X_1)} = \sigma_{X_1}$ and $\sqrt{\mathrm{var}(\theta)} = \sigma_\theta$.

Following the same argument as before, we keep $\Delta U(X_1^F,\rho,F)$ equal to 0 and implicit differentiation yields:

$$\frac{\partial X_1^F}{\partial \rho} = \left[\frac{-\dfrac{\partial \Delta U(\rho, X_1, F)}{\partial \rho}}{\dfrac{\partial \Delta U(\rho, X_1, F)}{\partial X_1}} \right]_{X_1 = X_1^F}$$

Now, $\dfrac{\partial \Delta U(\rho, X_1, F)}{\partial \rho} = -I\sigma_\theta \rho\left(1-\rho^2\right)^{-1/2}\left[C\Phi(C) + \varphi(C)\right] + I\sigma_\theta\left(1-\rho^2\right)^{1/2}\left[\dfrac{\partial C}{\partial \rho}\Phi(C)\right]$ and since

$$\frac{\partial C}{\partial \rho} = \frac{X_1 - \overline{X}_1}{\sigma_{X_1}\left(1-\rho^2\right)^{3/2}} + \frac{\rho\left[I\overline{\theta} - F\right]}{I\sigma_\theta\left(1-\rho^2\right)^{3/2}} \qquad \text{(A5),}$$

$$\frac{\partial \Delta U(\rho, X_1, F)}{\partial \rho} = -I\sigma_\theta \rho \left(1-\rho^2\right)^{-1/2} \frac{I\left(\bar{\theta} + \rho \frac{\sigma_\theta}{\sigma_{X_1}}\left(X_1 - \bar{X}_1\right)\right) - F}{I\sigma_\theta \left(1-\rho^2\right)^{1/2}} \Phi(C)$$

$$-I\sigma_\theta \rho \left(1-\rho^2\right)^{-1/2} \varphi(C) + \frac{I\sigma_\theta \left(X_1 - \bar{X}_1\right) + \rho\sigma_{X_1}\left[I\bar{\theta} - F\right]}{\sigma_{X_1}(1-\rho^2)} \Phi(C)$$

$$= \left[-\rho \frac{(I\bar{\theta} - F) + I\rho \frac{\sigma_\theta}{\sigma_{X_1}}\left(X_1 - \bar{X}_1\right)}{(1-\rho^2)} + \frac{I\sigma_\theta \left(X_1 - \bar{X}_1\right) + \rho\sigma_{X_1}\left[I\bar{\theta} - F\right]}{\sigma_{X_1}(1-\rho^2)} \right] \Phi(C) - I\sigma_\theta \rho \left(1-\rho^2\right)^{-1/2} \varphi(C)$$

$$= \frac{I\sigma_\theta \left(X_1 - \bar{X}_1\right)}{\sigma_{X_1}} \Phi(C) - \frac{I\sigma_\theta \rho}{\left(1-\rho^2\right)^{1/2}} \varphi(C) \qquad \text{(A6)}$$

It suffices now to use equations (A2) and (A6) to get:

$$\frac{\partial X_1^F}{\partial \rho} = -\frac{\frac{I\sigma_\theta \left(X_1^F - \bar{X}_1\right)}{\sigma_{X_1}} \Phi(C^{X_1^F}) - \frac{I\sigma_\theta \rho}{\left(1-\rho^2\right)^{1/2}} \varphi(C^{X_1^F})}{I\rho \frac{\sigma_\theta}{\sigma_{X_1}} \Phi(C^{X_1^F})} = \frac{\sigma_{X_1}}{\left(1-\rho^2\right)^{1/2}} \cdot \frac{\varphi(C^{X_1^F})}{\Phi(C^{X_1^F})} - \frac{(X_1^F - \bar{X}_1)}{\rho} \qquad \text{(A7).}$$

To prove that this expression is positive, we need to show that $X_1^F \leq \bar{X}_1$ for all values of F. For that purpose, we introduce the quantity X_1^* such that $E(\theta \mid X_1 = X_1^*) = r_2$: $X_1^* \equiv \bar{X}_1 + (r_2 - \bar{\theta}) \frac{\sigma_{X_1}}{\rho \sigma_\theta}$. Then $\frac{\partial X_1^*}{\partial \rho} = (\bar{\theta} - r_2) \frac{\sigma_{X_1}}{\sigma_\theta} \cdot \frac{1}{\rho^2} = \frac{\bar{X}_1 - X_1^*}{\rho}$ and our (realistic) assumption that $E(\theta) > r_2$ ensures that $X_1^* < \bar{X}_1$. Now, the last step is that, according to equation (3) in the text,

$$\frac{\partial U(Q, \rho, X_1^*, F)}{\partial Q} = \sqrt{\operatorname{var}(\theta \mid X_1^*)}\, \varphi\left(\frac{Ir_2 - F}{Q\sqrt{\operatorname{var}(\theta \mid X_1^*)}} \right) > 0$$

for all $Q \in [0,I]$ since $E(\theta \mid X_1 = X_1^*) = r_2$. Therefore $\Delta U(\rho, X_1^*, F) = U(I, \rho, X_1^*, F) - U(0, \rho, X_1^*, F) > 0$. Now $\Delta U(\rho, X_1^F, F_2) = 0$ by definition of X_1^F. The strict increasingness of ΔU in X_1 ensures that $X_1^* > X_1^F$ for all $F \in [0, Ir_2)$.

We have proved our result: $X_1^F < X_1^* < \bar{X}_1$ and the expression (A7) becomes obviously positive.

7. Proof of Theorem 2

We will prove here that the earlier uncertainty is resolved (i.e. the higher ρ), the lower the extent of risk shifting X_1^0-X_1^F. To do that, it suffices to show that $\partial(X_1^0-X_1^F)/\partial\rho<0$, i.e. $\partial X_1^0/\partial\rho<\partial X_1^F/\partial\rho$. One way to show that is to prove that $\partial^2 X_1^F/\partial\rho\partial F>0$.

Now, drawing from the results in the previous section and differentiating (A7) with respect to F,

$$\frac{\partial^2 X_1^F}{\partial\rho\partial F}=\frac{-\sigma_{X_1}C^{X_1^F}\varphi(C^{X_1^F})\left(\frac{\partial C^{X_1^F}}{\partial F}+\frac{\partial C^{X_1^F}}{\partial X_1^F}\cdot\frac{\partial X_1^F}{\partial F}\right)\Phi(C^{X_1^F})-\left(\frac{\partial C^{X_1^F}}{\partial F}+\frac{\partial C^{X_1^F}}{\partial X_1^F}\cdot\frac{\partial X_1^F}{\partial F}\right)\sigma_{X_1}\left[\varphi(C^{X_1^F})\right]^2}{(1-\rho^2)^{1/2}\left[\Phi(C^{X_1^F})\right]^2}-\frac{\partial X_1^F}{\partial F}\cdot\frac{1}{\rho}$$

$$=\frac{-\sigma_{X_1}\varphi(C^{X_1^F})D\left[C^{X_1^F}\Phi(C^{X_1^F})+\varphi(C^{X_1^F})\right]}{(1-\rho^2)^{1/2}\left[\Phi(C^{X_1^F})\right]^2}-\frac{\Phi(C^{X_1^F})-1}{\lambda\rho^2\frac{\sigma_\theta}{\sigma_{X_1}}\Phi(C^{X_1^F})}$$

where we took the expression for $\partial X_1^F/\partial F$ from (A4) and

$$D\equiv\left(\frac{\partial C^{X_1^F}}{\partial F}+\frac{\partial C^{X_1^F}}{\partial X_1^F}\cdot\frac{\partial X_1^F}{\partial F}\right)\left(\frac{-1}{\lambda\sigma_\theta(1-\rho^2)^{1/2}}+\frac{\rho}{\sigma_{X_1}(1-\rho^2)^{1/2}}\frac{\Phi(C^{X_1^F})-1}{\lambda\rho\frac{\sigma_\theta}{\sigma_{X_1}}\Phi(C^{X_1^F})}\right)=\frac{-1}{\lambda\sigma_\theta(1-\rho^2)^{1/2}\Phi(C^{X_1^F})}$$

Therefore

$$\frac{\partial^2 X_1^F}{\partial\rho\partial F}=\frac{\rho^2\varphi(C^{X_1^F})\left[C^{X_1^F}\Phi(C^{X_1^F})+\varphi(C^{X_1^F})\right]+(1-\rho^2)\left[1-\Phi(C^{X_1^F})\right]\left[\Phi(C^{X_1^F})\right]^2}{\rho^2\frac{\lambda\sigma_\theta}{\sigma_{X_1}}(1-\rho^2)\left[\Phi(C^{X_1^F})\right]^3}$$

It now suffices to remember that by definition of X_1^F, $\Delta U(\rho,X_1^F,F)=0$ or, from looking at (A1), $C^{X_1^F}\Phi(C^{X_1^F})+\varphi(C^{X_1^F})=\frac{\lambda r_2-F_2}{\lambda\sigma_\theta(1-\rho^2)^{1/2}}>0$. Hence, the expression above is obviously positive. QED.

Finally, when ρ tends to 0, X_1 does not give any information on θ and the managers of an all-equity firm and a levered firm will compare the unconditional expectation of θ with r_2. Since $E(\theta)>r_2$, they will always invest in the risky technology, regardless of the particular realization x_1: $\lim_{\rho\to0}X_1^0=\lim_{\rho\to0}X_1^F=-\infty$. However, from (A7) we infer that

$$\frac{\partial(X_1^0-X_1^F)}{\partial\rho}=\frac{\sigma_{X_1}}{(1-\rho^2)^{1/2}}\left(\frac{\varphi(C^{X_1^0})}{\Phi(C^{X_1^0})}-\frac{\varphi(C^{X_1^F})}{\Phi(C^{X_1^F})}\right)+\frac{X_1^F-X_1^0}{\rho}$$

so that $\lim_{\rho\to0}\partial(X_1^0-X_1^F)/\partial\rho=-\infty$ (for ρ close to 0, $C^{X_1^0}$ and $C^{X_1^F}$ tend to finite constants). This is enough to ensure that $\lim_{\rho\to0}(X_1^0-X_1^F)=+\infty$: the extent of risk shifting is discontinuous in $\rho=0$.

8. Computing the t=0 *Market* Value of the Firm

We will here, for the sake of generality, consider a firm with an investment policy of [ξ]: at t=1: if $x_1<\xi$, the manager invests everything in riskless bonds and $V_1(\rho)=I$; if $x_1\geq\xi$, everything gets invested in the risky technology θ so that the t=1 market value of the firm conditional on investment is:

$$V_1(\rho,x_1|x_1>\xi)=\frac{1}{r_2}\int_0^\infty I\theta\, dP(\theta|X_1)=\frac{1}{r_2}\left[IE(\theta\,|\,X_1)\Phi\left(\frac{E(\theta\,|\,X_1)}{\sigma_{\theta|X_1}}\right)+I\sigma_{\theta|X_1}\varphi\left(\frac{E(\theta\,|\,X_1)}{\sigma_{\theta|X_1}}\right)\right]$$

so that the t=0 market value of the firm is:

$$\begin{aligned}V_0(\rho,\xi)&=\frac{1}{r_1r_2}\left\{\int_{-\infty}^{\xi} Ir_2\,dP(X_1)+I\int_{\xi}^{\infty}\left[E(\theta\,|\,X_1)\Phi\left(\frac{E(\theta\,|\,X_1)}{\sigma_{\theta|X_1}}\right)+\sigma_{\theta|X_1}\varphi\left(\frac{E(\theta\,|\,X_1)}{\sigma_{\theta|X_1}}\right)\right]dP(X_1)\right\}\\&=\frac{I}{r_1r_2}\left\{r_2\Phi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)+\sigma_\theta(1-\rho^2)^{1/2}\int_{\xi}^{\infty}\left[A\Phi(A)+\varphi(A)\right]dP(X_1)\right\}\end{aligned}\qquad\text{(A8)}$$

where A stands for $E(\theta\,|\,X_1)/\sigma_{\theta|X_1}$ and is a function of a particular x_1 (note that it is nothing else than the expression C when F=0). The above is different from the *social* value of the firm, which is equal to:

$$\begin{aligned}V_0^S(\rho,\xi)&=\frac{1}{r_1r_2}\left\{\int_{-\infty}^{\xi} Ir_2\,dP(X_1)+\int_{\xi}^{+\infty} IE(\theta\,|\,X_1)dP(X_1)\right\}\\&=\frac{I}{r_1r_2}\left\{\bar{\theta}-(\bar{\theta}-r_2)\Phi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)+\rho\sigma_\theta\varphi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)\right\}\end{aligned}$$

The difference is accounted by the third negative claim, i.e. the responsibility to cover a negative realization of θ (e.g. environmental clean up costs) which we denote $G_0(\rho,\xi)$:

$$G_0(\rho,\xi)=\frac{I}{r_1r_2}\left\{\bar{\theta}\left[1-\Phi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)\right]+\rho\sigma_\theta\varphi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)-\sigma_\theta(1-\rho^2)^{1/2}\int_{\xi}^{\infty}\left[A\Phi(A)+\varphi(A)\right]dP(X_1)\right\}$$

The all-equity (resp. levered) firm will then substitute X_1^0 (resp. X_1^F) for ξ. Now,

$$\frac{\partial V_0(\rho,\xi)}{\partial\xi}=\frac{I}{r_1r_2}\left\{\frac{r_2}{\sigma_{X_1}}\varphi\left(\frac{\xi-\bar{X}_1}{\sigma_{X_1}}\right)-\sigma_\theta(1-\rho^2)^{1/2}\left[A^\xi\Phi(A^\xi)+\varphi(A^\xi)\right]\varphi_{X_1}(\xi)\right\}\qquad\text{(A9)}$$

where $\varphi_{X_1}(\xi)=\frac{1}{\sqrt{2\pi}\,\sigma_{X_1}}e^{-\frac{1}{2}\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)^2}$ denotes the probability density function of X_1 taken in $x_1=\xi$ and, as before, the superscript ξ on A means that this expression is also taken in $x_1=\xi$. We then have to remember that

- $\sigma_\theta\left(1-\rho^2\right)^{1/2}\left[A^{X_1^0}\Phi(A^{X_1^0})+\varphi(A^{X_1^0})\right]=r_2$ (when taken in $x_1=X_1^0$, the expressions A and C are equal);

- the functions $f: x \mapsto x\Phi(x)+\varphi(x)$ and $g: x \mapsto E(\theta \,|\, X = x)$ are increasing in x; hence $A^x \equiv \frac{E(\theta \,|\quad x)}{\sigma_{\ |X_1}} < \frac{E(\theta \,|\, X_1 = \ \ _1)}{\sigma_{\theta|X_1}} \quad A^{X_1^0}$ since we consider only cutoffs ξ's no larger than X_1^0 by Lemma 2; this in turn leads to the fact that $A^x\Phi(A^x)+\varphi(A^x) < A^{X_1^F}\Phi(A^{X_1^F})+\varphi(A^{X_1^F}) \equiv \frac{r_2}{\sigma_\theta(1-\rho^2)^{1/2}}$;
- and $\varphi_{X_1}(x) = \varphi\left(\frac{x-\bar{X}}{\sigma_{X_1}}\right)/\sigma_{X_1}$;

and the expression (A9) becomes strictly greater than zero: now, since by (A4) $\partial X_1^F/\partial F<0$, applying the chain rule ensures that $\partial V_0(\rho,X_1^F)/\partial F<0$: *agency costs are strictly increasing in the amount of risky debt outstanding*.

It is also interesting to note that the value of the firm (all-equity as well as levered) is increasing in ρ: from the value of the firm as given by (A8),

$$\frac{r_1r_2}{I}\cdot\frac{\partial V_0(\rho,x)}{\partial\rho} = \frac{r_2}{\sigma_{X_1}}\varphi\left(\frac{x-\bar{X}}{\sigma_{X_1}}\right)\frac{\partial x}{\partial\rho} - \sigma_\theta\rho\left(1-\rho^2\right)^{-1/2}\int_{X_1^F}^{\infty}\left[A\Phi(A)+\varphi(A)\right]dP(X_1)$$
$$+\sigma_\theta\left(1-\rho^2\right)^{1/2}\left\{\int_{x}^{\infty}\frac{\partial\left[A\Phi(A)+\varphi(A)\right]}{\partial\rho}dP(X_1) - \left[A^x\Phi(A^x)+\varphi(A^x)\right]\varphi_{X_1}(x)\frac{\partial x}{\partial\rho}\right\}$$

Since $A^x\Phi(A^x)+\varphi(A^x) < A^{X_1^F}\Phi(A^{X_1^F})+\varphi(A^{X_1^F}) \equiv r_2/(\sigma_\theta(1-\rho^2)^{1/2})$, the derivative above is no smaller than

$$-\sigma_\theta\rho\left(1-\rho^2\right)^{-1/2}\int_{x}^{\infty}\left[A\Phi(A)+\varphi(A)\right]dP(X_1) + \sigma_\theta\left(1-\rho^2\right)^{1/2}\int_{x}^{\infty}\Phi(A)\frac{\partial A}{\partial\rho}dP(X_1)$$
$$= \sigma_\theta\int_{x}^{\infty}\Phi(A)\frac{X_1-\bar{X}_1}{\sigma_{X_1}}dP(X_1) - \rho\,\sigma_\theta\left(1-\rho^2\right)^{-1/2}\int_{x}^{\infty}\varphi(A)dP(X_1) \qquad \text{(using (A5) with F = 0 for } \frac{\partial A}{\partial\rho}\text{)}$$
$$= \sigma_\theta\int_{\frac{x-\bar{X}_1}{\sigma_{X_1}}}^{\infty} z\Phi(a_A+b_Az)\varphi(z)dz - \rho\,\sigma_\theta\left(1-\rho^2\right)^{-1/2}\int_{\frac{x-\bar{X}_1}{\sigma_{X_1}}}^{\infty}\varphi(a_A+b_Az)\varphi(z)dz$$

where $a_A = \frac{\bar{\theta}}{\sigma_\theta(1-\rho^2)^{1/2}}$, and $b_A = \frac{\rho}{(1-\rho^2)^{1/2}}$ are constants and z is a standard normal variate. Using equations (15) and (26) from Carr and Rubinstein (1995) to integrate the above normals and simplifying, the above expression is equal to

$$\sigma_\theta\Phi(A^x)\varphi\left(\frac{x-\bar{X}}{\sigma_{X_1}}\right) > 0 \qquad \text{QED}$$

Firm value is thus bounded below by

$$V_0(X_1^0,0) = V_0(X_1^F,0) = IE(\theta\,|\,\theta \geq 0)/r_1r_2 = I\left[\bar{\theta}\,\Phi(\bar{\theta}/\sigma_\theta)+\sigma_\theta\varphi(\bar{\theta}/\sigma_\theta)\right]/r_1r_2 > I\bar{\theta}/r_1r_2\,.$$

The difference between this and I/r_1, the t=0 present value of investing I dollars at t=1, is the surplus of the entrepreneur when he sells the firm to shareholders and bondholders at t=0. Alternatively, it can be seen as the "option value" of the firm, i.e. the value of the opportunity to invest in a technology with $E(\theta)>r_2$.

9. Temporal Resolution of Uncertainty and Bond Yields.

We will here price the bonds of the firm in terms of an arbitrary X_1-cutoff ξ for the sake of generality (we'll use this formula later for an X_1-cutoff different from X_1^F when we consider partial contracting).

- As of t=1, if $x_1<\xi$, the manager will invest in riskless bills and the bond will be worth F at t=2 (since $Ir_2>F$); if $x_1\geq\xi$, the manager will invest in the risky technology and the bond will be worth $\max[0,\min(I\theta,F)]$ [58] at t=2. So conditional on investing in the riskless technology, the t=1 value of the bond is F/r_2. Conditional of investing in the risky technology, the value of the bond as of t=1, denoted $B_1(\rho,F,x_1)$, is:

$$
\begin{aligned}
B_1(\rho,F,x_1|x_1>\xi) &= \frac{1}{r_2}\left\{ I\int_0^{F/I} \theta\, dP(\theta|X_1) + F\int_{F/I}^{\infty} dP(\theta|X_1)\right\} \\
&= \frac{1}{r_2}\left\{ I\int_{\frac{-E(\theta|X_1)}{\sigma_{\theta|X_1}}}^{\frac{F/I-E(\theta|X_1)}{\sigma_{\theta|X_1}}} \left(E(\theta|X_1)+\sigma_{\theta|X_1} z\right) dP(z) + F\int_{\frac{F/I-E(\theta|X_1)}{\sigma_{\theta|X_1}}}^{\infty} dP(z)\right\} \\
&= \frac{1}{r_2}\left\{ F\Phi(C) + IE(\theta|X_1)\left[\Phi(A)-\Phi(C)\right] + I\sigma_{\theta|X_1}\left[\varphi(A)-\varphi(C)\right]\right\} \\
&= \frac{I\sigma_{\theta|X_1}}{r_2}\left[A\Phi(A)+\varphi(A)-C\Phi(C)-\varphi(C)\right]
\end{aligned}
$$

where as before $A=\dfrac{E(\theta|X_1=x_1)}{\sigma_{\theta|X_1}}$ and $C=A-F/I\sigma_{\theta|X_1}$; once more, it is obvious from the derivation that $\Phi(C)$, which is a function of the particular realization x_1 that occurred, is nothing else than the t=1 probability of being solvent and that $\Phi(A)$ is the probability of θ being positive.

NB: in the same line of thought as before, note that the next-to-last line in the expression of $B_1(\rho,F,x_1)$ tells us the following: at t=1, bondholders expect the firm to be solvent (and to receive F) with probability $\Phi(C)$ and believe that with a probability of $\Phi(A)$-$\Phi(C)$ the firm will not be solvent but still have positive

[58] Since the return on the risky technology is normally distributed, it can assume an arbitrarily negative value. As stated in the text, we consider in that case that the society at large has to bear the consequences of a negative realization of θ (product liability suit, environmental catastrophe etc.). However, we can choose the parameters so as to make the

value, in which case they will take over, receiving on average $\mathrm{IE}(\theta|X_1=x_1)$; the last term is a convexity adjustment.

- So as of t=0, the bond price is equal to:

$$B_0(\rho,F,x)=\frac{1}{r_1r_2}\left\{F\Phi\left(\frac{x-\bar{X}}{\sigma_{X_1}}\right)+I\sigma_\theta(1-\rho^2)^{1/2}\int_x^\infty\left[A\Phi(A)+\varphi(A)-C\Phi(C)-\varphi(C)\right]dP(X_1)\right\} \qquad \text{(A10)}$$

as stated in the text.

We are more interested, however, in how the yield premium evolves with ρ, the pattern of temporal resolution of uncertainty. Now, since the debt has the form of a zero-coupon bond, we define the per period yield premium demanded on the debt as $\sqrt{\frac{F}{r_1r_2B_0(\rho,F)}}$ -1 (we remind the reader here that the bond has a maturity of two periods). Proving that this yield premium is decreasing in ρ is therefore equivalent to proving that bond prices are increasing in ρ. Then, from (A10),

$$\begin{aligned}\frac{\partial r_1r_2B_0(\rho,F,x)}{\partial\rho}=&\frac{F}{\sigma_{X_1}}\varphi\left(\frac{x-\bar{X}_1}{\sigma_{X_1}}\right)\frac{\partial x}{\partial\rho}-I\sigma_\theta\rho\left(1-\rho^2\right)^{-1/2}\int_x^\infty\left[A\Phi(A)+\varphi(A)-C\Phi(C)-\varphi(C)\right]dP(X_1)\\&+I\sigma_\theta\left(1-\rho^2\right)^{1/2}\left\{\int_x^\infty\frac{\partial\left[A\Phi(A)+\varphi(A)\right]}{\partial\rho}dP(X_1)-\left[A^x\Phi(A^x)+\varphi(A^x)\right]\varphi_{X_1}(x)\frac{\partial x}{\partial\rho}\right\}\\&-I\sigma_\theta\left(1-\rho^2\right)^{1/2}\left\{\int_x^\infty\frac{\partial\left[C\Phi(C)+\varphi(C)\right]}{\partial\rho}dP(X_1)-\left[C^x\Phi(C^x)+\varphi(C^x)\right]\varphi_{X_1}(x)\frac{\partial x}{\partial\rho}\right\}\end{aligned} \qquad \text{(A11)}$$

Here, we need to come back to our particular case where the cutoff is X_1^F. We then have to remember that

- $I\sigma_\theta\left(1-\rho^2\right)^{1/2}\left[C^{X_1^F}\Phi(C^{X_1^F})+\varphi(C^{X_1^F})\right]\equiv Ir_2-F$;
- $A^{X_1^F}\Phi(A^{X_1^F})+\varphi(A^{X_1^F})<A^{X_1^0}\Phi(A^{X_1^0})+\varphi(A^{X_1^0})\equiv\frac{r_2}{\sigma_\theta(1-\rho^2)^{1/2}}$;
- $\varphi_{X_1}(X_1^F)=\varphi\left(\frac{X_1^F-\bar{X}}{\sigma_{X_1}}\right)/\sigma_{X_1}$; and $\partial X_1^F/\partial\rho>0$.

and the expression above becomes strictly greater than:

---

probability of a negative cash flow $\left(\int_{X_1^F}^\infty\left(\int_{-\infty}^0 dP(\theta\mid X_1)\right)dP(X_1)\right)$ arbitrarily small. With our parameters ($E(\theta)=1.15$, $\sigma_\theta=0.6$), it is never larger than 2.76% for any value of ρ.

$$-I\sigma_\theta\left(1-\rho^2\right)^{-1/2}\left\{\int_{X_1^F}^{\infty}\left[A\Phi(A)+\rho\varphi(A)-\left(1-\rho^2\right)\Phi(A)\frac{\partial A}{\partial\rho}\right]dP(X_1)-\int_{X_1^F}^{\infty}\left[C\Phi(C)+\rho\varphi(C)-\left(1-\rho^2\right)\Phi(C)\frac{\partial C}{\partial\rho}\right]dP(X_1)\right\}$$

$$=I\rho\sigma_\theta\left(1-\rho^2\right)^{-1/2}\int_{X_1^F}^{\infty}\left[\varphi(C)-\varphi(A)\right]dP(X_1)-I\sigma_\theta\int_{X_1^F}^{\infty}\frac{(X_1-\overline{X}_1)}{\sigma_{X_1}}\left[\Phi(C)-\Phi(A)\right]dP(X_1)$$

$$=I\rho\sigma_\theta\left(1-\rho^2\right)^{-1/2}\int_{\frac{X_1^F-\overline{X}_1}{\sigma_{X_1}}}^{\infty}\left[\varphi(a_C+b_Cz)-\varphi(a_A+b_Az)\right]\varphi(z)dz-I\sigma_\theta\int_{\frac{X_1^F-\overline{X}_1}{\sigma_{X_1}}}^{\infty}\left[\Phi(a_C+b_Cz)-\Phi(a_A+b_Az)\right]z\varphi(z)dz$$

where, as in the derivation of $\partial V_0(\rho)/\partial\rho$, we used the expression (A5) for $\partial C/\partial\rho$ and $\partial A/\partial\rho$ and $a_A=\frac{\overline{\theta}}{\sigma_\theta(1-\rho^2)^{1/2}}$, $a_C=a_A-\frac{F/I}{\sigma_\theta(1-\rho^2)^{1/2}}$, $b_A=b_C=\frac{\rho}{(1-\rho^2)^{1/2}}$ are all constants and z is a standard normal variate. Using equations (15) and (26) from Carr and Rubinstein (1995) to integrate the above normals and simplifying, the above expression is equal to

$$I\sigma_\theta\varphi\left(\frac{X_1^F-\overline{X}_1}{\sigma_{X_1}}\right)\left[\Phi(A^{X_1^F})-\Phi(C^{X_1^F})\right]$$

which is positive since for any x_1 (in particular X_1^F), $C=A-F/I\sigma_{\theta|X_1}<A$. QED.

It is also of interest to notice the continuity of bond prices when ρ tends to 1: the price of the debt tends to F/r_1r_2 (the yield premium tends to 0): as we saw earlier in the text, when x_1 reveals θ, the manager of a levered firm adopts the same cutoff policy as the manager of an all-equity firm: the cutoff X_1^F is such that $E(\theta|X_1^F)=r_2$ and since $Ir_2>F$, C tends to infinity for all x_1 greater than X_1^F. Hence $C\Phi(C)+\varphi(C)\sim C$, $A\Phi(A)+\varphi(A)\sim A$ and

$$\lim_{\rho\to1}\left\{I\sigma_{\theta|X_1}\int_{X_1^F}^{\infty}\left[A\Phi(A)+\varphi(A)-C\Phi(C)+\varphi(C)\right]dP(X_1)\right\}=\lim_{\rho\to1}\left\{I\sigma_\theta\left(1-\rho^2\right)^{1/2}\int_{X_1^F}^{\infty}\left[\frac{E(\theta|X_1)}{\sigma_\theta\left(1-\rho^2\right)^{1/2}}-\frac{E(\theta|X_1)-F/I}{\sigma_\theta\left(1-\rho^2\right)^{1/2}}\right]dP(X_1)\right\}$$

$$=\int_{X_1^F}^{\infty}F\,dP(X_1)=F\left[1-\Phi\left(\frac{X_1^F-\overline{X}_1}{\sigma_{X_1}}\right)\right]$$

It now suffices to replace the above in equation (A10) to obtain the desired result: $\lim_{\rho\to1}B_0(\rho,F,X_1^F)=\frac{F}{r_1r_2}$.

The intuition is as explained in the text: when x_1 reveals θ, the firm is always solvent and the debt is riskless.

Finally, in much the same way as we did for bond prices, it can be shown that share price as well as the value of the claim held (unwillingly) by taxpayers are increasing in ρ: an earlier resolution of uncertainty is Pareto-improving.

10. Mitigating Risk-Shifting through Partial Contracting

Our first task is to prove Proposition 4, i.e. $\frac{\partial Y_1^0}{\partial r_{XY}} > 0$. This is done in much the same way as in Appendix 6 and we'll go quickly through the proof. Keeping ÄU(Y$_1$,ρ,0) constant at 0 (this is the implicit definition of Y_1^0, using the same notation as earlier), total differentiation yields:

$$\frac{\partial Y_1^0}{\partial r_{XY}} = \left. \frac{-\dfrac{\partial \Delta U(r, Y_1, 0)}{\partial r_{XY}}}{\dfrac{\partial \Delta U(r, X_1, 0)}{\partial Y_1}} \right|_{Y_1 = Y_1^0}$$

Now, $\frac{\partial \Delta U(Y_1, r, 0)}{\partial Y_1} = I r_m r_{XY}^2 \frac{s_q}{s_{X_1}} \Phi(A^{Y_1})$, where A^{Y_1} denotes the expression $\frac{E(q \mid Y_1)}{s_{q|Y_1}}$ and we used the fact that E(θ|Y$_1$) depends on ρ$_{XY}$ both through $r_b = r_{XY} r_m$ and through $s_{Y_1} = \frac{s_{X_1}}{r_{XY}}$.

Finally, since $\frac{\partial A^{Y_1}}{\partial r_{XY}} = \frac{r_m r_{XY}(Y_1 - \overline{X}_1)(2 - r_m^2 r_{XY}^2) + r_m^2 r_{XY} \overline{q}}{s_q \left(1 - r_m^2 r_{XY}^2\right)^{3/2}}$,

$$\frac{\partial \Delta U(Y_1, r, 0)}{\partial r_{XY}} = -I s_q r_m^2 r_{XY} (1 - r_m^2 r_{XY}^2)^{-1/2} \left[A^{Y_1} \Phi(A^{Y_1}) + j(A^{Y_1}) \right] + I s_q (1 - r_m^2 r_{XY}^2)^{1/2} \Phi(A^{Y_1}) \frac{\partial A^{Y_1}}{\partial r_{XY}}$$

$$= 2 I r_m r_{XY} \frac{s_q}{s_{X_1}} (Y_1 - \overline{X}_1) \Phi(A^{Y_1}) - I s_q r_m^2 r_{XY} (1 - r_m^2 r_{XY}^2)^{-1/2} j(A^{Y_1})$$

and $\frac{\partial Y_1^0}{\partial r_{XY}} = \frac{r_1 j(A^{Y_1})}{\frac{r_{XY}}{s_{X_1}} (1 - r_m^2 r_{XY}^2)^{1/2} \Phi(A^{Y_1})} - 2 \frac{Y_1^0 - \overline{X}_1}{r_{XY}}$, which is positive since $Y_1^0 < X_1^0$ as indicated in the text and $X_1^0 < \overline{X}_1$ (we introduced the quantity X_1^* in Appendix 6 and proved that $X_1^F < X_1^* < \overline{X}_1$ for all values of F∈[0,Ir$_2$), in particular F=0). Finally, the increasingness of the cutoff in ρ and its decreasingness in the variance of the signal (i.e. s_{X_1} for X_1^0, s_{Y_1} for Y_1^0), coupled with the fact that ρ$_b$=ρ$_m$ρ$_{XY}$ and $s_{Y_1} = s_{X_1} / r_{XY}$ ensures that $\lim_{r_{XY} \to 0} Y_1^0 < \lim_{r_m \to 0} X_1^0 < \lim_{r_m \to 0} X_1^* = -\infty$. It should be obvious that $\lim_{r_{XY} \to 1} Y_1^0 = X_1^0$ for any ρ$_m$.

The next step is to prove Theorem 4. As we saw in the text, the new cutoff will be $\max(X_1^F, Y_1^0)$. Hence the benefit from contracting is $\max(Y_1^0 - X_1^F, 0)$. Using (A7), *mutatis mutantis*, we come to the conclusion that $\frac{\partial Y_1^0}{\partial r_b} = \frac{s_{Y_1}}{\left(1 - r_b^2\right)^{1/2}} \cdot \frac{j(A^{Y_1^0})}{\Phi(A^{Y_1^0})} - \frac{(Y_1^0 - \overline{X}_1)}{r_b}$ where $A^{Y_1^0}$ denotes the expression $\frac{E(q \mid y_1 = Y_1^0)}{s_{q|Y_1}}$, so that

$\frac{\partial Y_1^0}{\partial r_m} = \frac{\partial Y_1^0}{\partial r_b} \frac{\partial r_b}{\partial r_m} = \frac{s_{X_1}}{\left(1 - r_m^2 r_{XY}^2\right)^{1/2}} \cdot \frac{j(A^{Y_1^0})}{\Phi(A^{Y_1^0})} - \frac{(Y_1^0 - \overline{X}_1)}{r_m} > 0$ and

$$\frac{\partial(Y_1^0 - X_1^F)}{\partial \rho_m} = \frac{\sigma_{X_1}}{\left(1-\rho_m^2\rho_{XY}^2\right)^{1/2}}\cdot\frac{\varphi(A^{Y_1^0})}{\Phi(A^{Y_1^0})} - \frac{\sigma_{X_1}}{\left(1-\rho_m^2\right)^{1/2}}\cdot\frac{\varphi(C^{X_1^F})}{\Phi(C^{X_1^F})} + \frac{(X_1^F - Y_1^0)}{\rho_m}$$

$$\leq \frac{\sigma_{X_1}}{\left(1-\rho_m^2\rho_{XY}^2\right)^{1/2}}\left[\frac{\varphi(A^{Y_1^0})}{\Phi(A^{Y_1^0})} - \frac{\varphi(C^{X_1^F})}{\Phi(C^{X_1^F})}\right] \text{when } Y_1^0 \geq X_1^F$$

Now, since the function $h: x \mapsto \varphi(x)/\Phi(x)$ is decreasing, it suffices to show that $A^{Y_1^0} > C^{X_1^F}$ to prove the negativity of $\partial(Y_1^0 - X_1^F)/\partial\rho_m$ when $Y_1^0 \geq X_1^F$. This is rather easy and left to the reader[59].

Now, when ρ_{XY} tends to 1, Y_1^0 tends to X_1^0 and $\lim_{\rho_m \to 0}\left(Y_1^0 - X_1^F\right) = \lim_{\rho_m \to 0}\left(X_1^0 - X_1^F\right) = \infty$ (see Appendix 7), while when ρ_{XY} is arbitrarily small, $Y_1^0 - X_1^F$ is arbitrarily negative for any value of ρ_m, in particular in the neighborhhod of $\rho_m=0$. Given the fact that $\frac{\partial(Y_1^0 - X_1^F)}{\partial \rho_{XY}} > 0$ (see Proposition 4 in the text), there is a critical $\underline{\rho}_{XY}$ such that for all $\rho_{XY} < \underline{\rho}_{XY}$, $\lim_{\rho_m \to 0}\left(Y_1^0 - X_1^F\right) < 0$. Given that $\frac{\partial(Y_1^0 - X_1^F)}{\partial \rho_m} < 0$ (see previous paragraph), this ensures that for $\rho_{XY} < \underline{\rho}_{XY}$, Y_1^0 is *never greater than* X_1^F *for any* ρ_m. In other words, if Y_1^0 is not greater than X_1^F at low ρ_m values, it will not either for larger values of ρ_m. Now, let us look at the case where $\rho_{XY} \geq \underline{\rho}_{XY}$ (i.e. it is not always the case that $Y_1^0 < X_1^F$). When both ρ_m and ρ_{XY} tend to 1, $\lim_{\substack{\rho_m \to 1 \\ \rho_{XY} \to 1}}\left(Y_1^0 - X_1^F\right) = \lim_{\rho_m \to 1}\left(X_1^0 - X_1^F\right) = 0$ and the increasingness of Y_1^0 in ρ_{XY} ensures that there is no ρ_{XY} for which Y_1^0 is greater than X_1^F when ρ_m is close enough to 1 (however Y_1^0 might well be strictly less than X_1^F close to $\rho_m=1$ since $\rho_{XY}=1$ is still the ρ_{XY} that maximizes Y_1^0). But since $\rho_{XY} \geq \underline{\rho}_{XY}$, $\lim_{\rho_m \to 0}\left(Y_1^0 - X_1^F\right) = \infty$ (as we saw before). This, combined with the facts that $\lim_{\substack{\rho_m \to 1 \\ \rho_{XY} \to 1}}\left(Y_1^0 - X_1^F\right) = 0$, $\frac{\partial(Y_1^0 - X_1^F)}{\partial \rho_{XY}} > 0$ and $\frac{\partial(Y_1^0 - X_1^F)}{\partial \rho_m} < 0$ and using the Intermediate Value Theorem, ensures that there indeed exists a unique ρ_m^* such that, for a given ρ_{XY} greater than $\underline{\rho}_{XY}$, $Y_1^0 > X_1^F$ for $\rho_m < \rho_m^*$ and $Y_1^0 < X_1^F$ for $\rho_m > \rho_m^*$. For $\rho_{XY} < \underline{\rho}_{XY}$, we'll set $\rho_m^* = 0$. On $[0, \rho_m^*]$, the improvement Y_1^0-X_1^F in terms of risk-shifting is decreasing in ρ_m, and on $[\rho_m^*, 1]$ the improvement is constant at 0.

---

[59] Attention: we stated earlier that $C = A - F/(I\sigma_q(1-\rho^2)^{1/2})$. However, here $A^{Y_1^0}$ uses σ_Y and $\rho_b=\rho_m\rho_{XY}$, while $C^{X_1^F}$ uses σ_X and ρ_m.

Finally, we show that this ρ_m^* is increasing in ρ_{XY}: an implicit differentiation at $\rho_m = \rho_m^*$ yields, as $Y_1^0 - X_1^F$ is kept constant at 0:

$$\frac{\partial \rho_m^*}{\partial \rho_{XY}} = -\left.\frac{\partial Y_1^0 / \partial \rho_{XY}}{\partial (Y_1^0 - X_1^F)/\partial \rho_m}\right|_{\rho_m=\rho_m^*} > 0$$

since the numerator is positive as stated in Proposition 4 in the text and the denominator is negative as was just proven. Since $\rho_{XY} = \frac{\sigma_{X_1}}{\sqrt{\sigma_{X_1}^2 + \sigma_\varepsilon^2}}$, ρ_m^* is decreasing in σ_ε.

11. Partial Contracting, Temporal Resolution of Uncertainty and Bond Yields

- For a given ρ_{XY}, if $Y_1^0 \le X_1^F$ (i.e. for all ρ_m if $\rho_{XY} \le \underline{\rho}_{XY}$ or for $\rho_m \ge \rho_m^*$ otherwise), nothing is changed and we refer the reader to Appendix 9 for a proof that bond yields are decreasing in ρ_m.
- if $\rho_m < \rho_m^*$, shareholders will invest, on average[60], for all $x_1 > Y_1^0$. This yields a bond price similar to the one in (A10), with Y_1^0 used to replace the generic cutoff ξ. This in turn yields the expression (A11) for the first derivative of bond prices with respect to ρ_m, but with Y_1^0 and ρ_m in lieu of X_1^F and ρ (the reader should also be aware that in this derivative, $A^{Y_1^0}$ denotes $\frac{E(\theta \mid x_1 = Y_1^0)}{\sigma_{\theta|X_1}} = \frac{\bar{\theta} + \rho_m \sigma_\theta (Y_1^0 - \bar{X}_1)/\sigma_{X_1}}{\sigma_\theta (1-\rho_m^2)^{1/2}}$, as opposed to Appendix 10, where it denoted $\frac{E(\theta \mid y_1 = Y_1^0)}{\sigma_{\theta|Y_1}} = \frac{\bar{\theta} + \rho_b \sigma_\theta (Y_1^0 - \bar{X}_1)/\sigma_{Y_1}}{\sigma_\theta (1-\rho_b^2)^{1/2}}$). Since we consider here the case where $Y_1^0 > X_1^F$, $A^{Y_1^0}\Phi(A^{Y_1^0}) + \varphi(A^{Y_1^0}) > A^{X_1^0}\Phi(A^{X_1^0}) + \varphi(A^{X_1^0}) \equiv \frac{r_2}{\sigma_\theta (1-\rho^2)^{1/2}}$ and since $X_1^0 > Y_1^0$, $C^{Y_1^0}\Phi(C^{Y_1^0}) + \varphi(C^{Y_1^0}) > C^{X_1^F}\Phi(C^{X_1^F}) + \varphi(C^{X_1^F}) \equiv \frac{Ir_2 - F}{I\sigma_\theta (1-\rho^2)^{1/2}}$; this, coupled with the fact that $\partial Y_1^0 / \partial \rho_m > 0$ (see previous Appendix) is enough to show, through the exact same steps as in Appendix 9, that $\frac{\partial r_1 r_2 B_0(\rho, F, Y_1^0)}{\partial \rho} > I\sigma_\theta \varphi\left(\frac{Y_1^0 - \bar{X}}{\sigma_{X_1}}\right)\left[\Phi(A^{Y_1^0}) - \Phi(C^{Y_1^0})\right] > 0$. Hence bond prices are still increasing in ρ_m after optimal contracting; equivalently, the yield premium demanded is lower the quicker uncertainty is resolved. QED.

---

[60] As we noted in the text, shareholders will really invest for all realizations of X_1 greater than $Y_1^0 - \varepsilon_1$. But since bondholders do not observe ε_1, they will price securities using $E(\varepsilon)=0$ for ε_1.

It can also be shown that this yield premium is non-increasing in ρ_{XY}, the accuracy of the information Y_1 available to outsiders. As before, we will do that through looking at comparative statics involving bond prices:

- if $Y_1^0 \le X_1^F$ (i.e. $\rho_m \ge \rho_m^*$), $\dfrac{\partial B_0(\rho,F,X_1^F)}{\partial \rho_{XY}} = 0$;
- if $Y_1^0 > X_1^F$ (i.e. $\rho_m < \rho_m^*$),

$$\frac{\partial r_1 r_2 B_0(\rho,F,Y_1^0)}{\partial \rho_{XY}} = \frac{F}{\sigma_{X_1}}\varphi\left(\frac{Y_1^0-\bar{X}_1}{\sigma_{X_1}}\right) + I\sigma_\theta\left(1-\rho_m^2\right)^{1/2}\left[C^{Y_1^0}\Phi(C^{Y_1^0})+\varphi(C^{Y_1^0}) - A^{Y_1^0}\Phi(A^{Y_1^0})+\varphi(A^{Y_1^0})\right]\varphi_{X_1}(Y_1^0)\frac{\partial Y_1^0}{\partial \rho_{XY}}$$

Now, as noted before, $I\sigma_\theta(1-\rho_m^2)^{1/2}\left[C^{Y_1^0}\Phi(C^{Y_1^0})+\varphi(C^{Y_1^0})\right] > Ir_2 - F$ since $Y_1^0 > X_1^F$ and $I\sigma_\theta(1-\rho_m^2)^{1/2}\left[A^{Y_1^0}\Phi(A^{Y_1^0})+\varphi(A^{Y_1^0})\right] < Ir_2$ since $Y_1^0 < X_1^0$; this, combined with the positivity of $\partial Y_1^0/\partial \rho_{XY}$, yields the desired result: bond yields are decreasing in ρ_{XY} on $[0,\rho_m^*]$ and independent of ρ_{XY} on $[\rho_m^*,1]$. QED.

This effect is reinforced by the increasingness of ρ_m^* in ρ_{XY} (see previous Appendix): as the latter increases, the region where bond yields are strictly decreasing in ρ_{XY} widens.

Finally, the price of the equity $S_0(\rho,\xi)$ can be expressed as the difference between the market value of the firm (A8) and the price of bonds (A10):

$$S_0(\rho,x) = \frac{1}{r_1 r_2}\left\{(Ir_2-F)\Phi\left[\frac{x-\bar{X}}{\sigma_{X_1}}\right] + I\sigma_\theta(1-\rho_m^2)^{1/2}\int_x^\infty \left[C\Phi(C)+\varphi(C)\right]dP(X_1)\right\}$$

where the cutoff ξ used is $\max(X_1^F, Y_1^0)$. Now, if $X_1^F > Y_1^0$, share prices are insensitive to a change in ρ_{XY} until ρ_{XY} is large enough for Y_1^0 to exceed X_1^F, in which case they become sensitive to a further increase in ρ_{XY}. Then

$$\frac{\partial r_1 r_2 S_0(\rho,Y_1^0)}{\partial \rho_{XY}} = \frac{1}{\sigma_{X_1}}\varphi\left[\frac{Y_1^0-\bar{X}_1}{\sigma_{X_1}}\right]\left[Ir_2-F\right]\frac{\partial Y_1^0}{\partial \rho_{XY}} - I\sigma_\theta\left(1-\rho_m^2\right)^{1/2}\left[C^{Y_1^0}\Phi(C^{Y_1^0})+\varphi(C^{Y_1^0})\right]\varphi_{X_1}(Y_1^0)\frac{\partial Y_1^0}{\partial \rho_{XY}}$$

Given that $C^{Y_1^0}\Phi(C^{Y_1^0})+\varphi(C^{Y_1^0}) \underset{Y_1^0>X_1^F}{>} \dfrac{Ir_2-F}{I\sigma_\theta(1-\rho_m^2)^{1/2}}$ as noted earlier, this last derivative is negative.

References

Akerlof, George A., 1970, “The Market for ‘Lemons’: Qualitative Uncertainty and the Market Mechanism”, *Quarterly Journal of Economics*, pp. 488-500.

Amihud, Yakov, Kenneth Garbade and Marcel Kahan, "A New Governance Structure for Corporate Bonds", *Stanford Law Review*, vol. 51, February, pp. 447-492.

Black, Fisher and Myron Scholes, 1973, “The Pricing of Options and Corporate Liabilities”, *Journal of Political Economy* 81, 637-659.

Bradley, Michael, Gregg A. Jarrell and Han Kim, 1984, “On the Existence of an Optimal Capital Structure: Theory and Evidence”, *Journal of Finance*, July, pp.857-880.

Brennan, Michael J., 1995, "Corporate Finance Over the Past 25 Years", *Financial Management*, vol. 24, no. 2, pp. 9-22.

Brito, Jose A., 1999, “Risk-Avoidance Induced by Risky Debt”, Working Paper, Stern School of Business, New York University.

Bronte, S., 1982, “Japanese Finance: Markets and Institutions”, London: Euromoney Publications Limited.

DeAngelo, Harry and Ronald W. Masulis, 1980, “Optimal Capital Structure Under Corporate and Personal Taxation”, *Journal of Financial Economics*, March, pp. 3-29.

Degeorge, François, Jayendu Patel and Richard Zeckhauser, 1999, “Earnings Management to Exceed Thresholds”, *Journal of Business*, vol. 72 no. 1, pp.1-33.

Duffie, Darrell and David Lando, 1998, “Term Structure of Credit Spreads with Incomplete Accounting Information”, Working Paper, Stanford University.

Epstein, Larry G., 1980, “Decision Making and the Temporal Resolution of Uncertainty”, *International Economic Review*, vol. 21, no. 2, pp. 269-283.

Epstein, Larry G. and Stuart M. Turnbull, 1980, “Capital Asset Prices and the Temporal Resolution of Uncertainty”, *Journal of Finance*, vol. 35, no. 3, pp. 627-643.

Feller, W., 1976, “An Introduction to Probability Theory and its Applications”, vol. 2, New York: John Wiley and Sons.

Gavish, Bezalel and Avner Kalay, 1983, “On the Asset Substitution Problem”, *Journal of Financial and Quantitative Analysis*, vol. 18, no. 1, pp. 21-30.

Goldberger, A. S., 1964, “Econometric Theory”, New York: John Wiley and Sons.

Goswami, Gautam, Thomas Noe and Michael Rebello, 1995, “Debt financing under Asymmetric Information”, *Journal of Finance*, vol. L., no. 2, pp. 633-659, June.

Green, Richard C., 1984, “Investment Incentives, Debt and Warrants”, *Journal of Financial Economics*, vol. 13, pp. 115-136.

Guedes, Jose and Tim Opler, 1996, “The Determinants of the Maturity Structure of Corporate Debt Issues”, *Journal of Finance*, vol. 51, no. 1, December.

Hull, John C., 1993, “Options, Futures and Other Derivative Securities”, Prentice Hall, second edition.

Jensen, Michael C. and William H. Meckling, 1976, “Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure”, *Journal of Financial Economics*, vol. 3, pp. 305-360.

John, Kose, 1987, “Risk-Shifting Incentives and Signalling Through Corporate Capital Structure”, *Journal of Finance*, vol. 42, nº. 3, pp. 623-641, July.

John, Kose and N. K. Chidambaran, 1998, “Relationship Investing: Large Shareholder Monitoring with Managerial Cooperation”, Working Paper, Stern School of Business, New York University.

John, Teresa and Kose John, 1993, “Top-Management Compensation and Capital Structure”, *Journal of Finance*, vol. 48, no. 3, pp. 949-974, July.

John, Kose and Joshua Ronen, 1990, “Information Structures, Optimal Contracts and the Theory of the Firm”, *Journal of Accounting, Auditing & Finance 5*, pp. 61-95.

Kester, Carl W., 1986, “Capital and Ownership Structure: A Comparison of United States and Japanese Manufacturing Corporations”, *Financial Management*, pp. 5-16.

Leland, Hayne E., 1994, “Corporate Debt Value, Bond Covenants and Optimal Capital Structure”, *Journal of Finance* 49, pp. 1213-1252.

Leland, Hayne E., 1998, “Agency Costs, Risk Management and Capital Structure”, *Journal of Finance* 53 n^{o} 4, pp. 1213-1243.

Long, Michael S. and Ileen B. Malitz, 1986, "Investment Patterns and Financial Leverage", NBER Study edited by Benjamin Friedman, University of Chicago Press, pp.325-348.

Modigliani, Franco and Merton H. Miller, 1958, “The Cost of Capital, Corporation Finance and the Theory of Investment”, *American Economic Review* 48, pp. 261-297.

Myers, Stewart C., 1977, “Determinants of Corporate Borrowing”, *Journal of Financial Economics*, vol. 5, pp. 147-175.

Nabar, Prafulla, Richard C. Stapleton and Marti G. Subrahmanyam, 1988, “Default Risk, Resolution of Uncertainty and the Interest Rate on Corporate Loans”, *Studies in Banking and Finance* 5, pp. 221-245.

Reisz, Alexander, 1999, “Temporal Resolution of Uncertainty and Corporate Debt Yields: an Empirical Investigation”, Working Paper, Stern School of Business, New York University.

Robichek, A. A. and Stuart C. Myers, 1966, “Valuation of the Firm: Effects of Uncertainty in a Market Context”, *Journal of Finance*, vol. 21.

Ross, Stephen, 1978, “A Simple Approach to Valuation of Risky Streams”, *Journal of Business* 51, pp. 453-475.

Ross, Stephen, 1989, “Information and Volatility: the No-arbitrage Martingale Approach to Timing and Resolution Irrelevancy”, *Journal of Finance*, vol. XLIV, no. 1, March, pp. 1-17.

Rothschild, Michael and Joseph Stiglitz, 1970, “Increasing Risk I: A Definition”, *Journal of Economic Theory* 2, pp. 225-243.

Rothschild, Michael and Joseph Stiglitz, 1971, “Increasing Risk II: Its Economic Consequences”, *Journal of Economic Theory* 3, pp. 66-84.

Rubinstein, Mark and Peter Carr, 1995, “Appendix: Useful Facts about Normal Distributions”, in Rubinstein, Mark and Eric Reiner, “Exotic Options”, Working Paper, Haas School of Business, University of California at Berkeley.

Schwartz, E. and J. R. Aronson, 1967, “Some Surrogate Evidence in Support of the Concept of Optimal Capital Structure”, *Journal of Finance*, March, pp. 10-18.

Smith, Clifford W. Jr. and René Stulz, 1985, “The Determinants of Firm’s Hedging Policies”, *Journal of Financial and Quantitative Analysis* 20, pp. 391-405.

Smith, Clifford W. and Jerold B. Warner, 1979, “On Financial Contracting: an Analysis of Bond Covenants”, *Journal of Financial Economics*, vol. 7, pp. 117-161.

Stapleton, Richard C. and Marti G. Subrahmanyam, 1978, “A Multiperiod Equilibrium Asset Pricing Model”, *Econometrica*, vol. 46, no. 5, pp. 1077-1096.

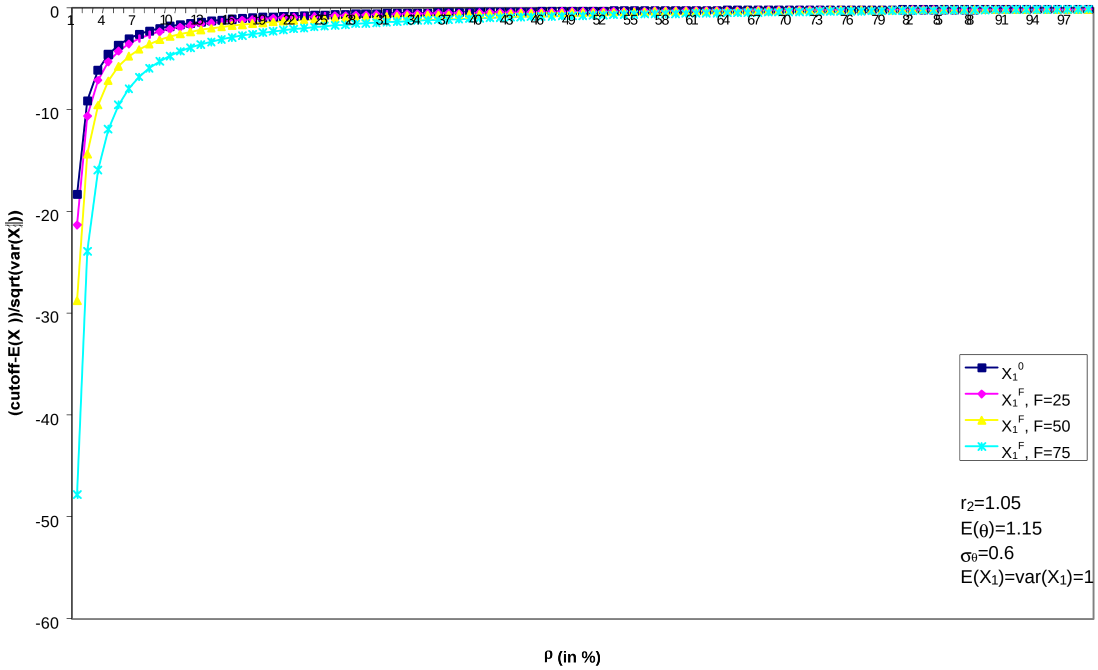
Stapleton, Richard C. and Marti G. Subrahmanyam, 1984, “The Valuation of Multivariate Contingent Claims in Discrete Time Models”, *Journal of Finance*, vol. 39,no.3, pp. 207-228, March.

Stohs, Mark Hoven and David C. Mauer, 1996, “The Determinants of Corporate Debt Maturity Structure”, *Journal of Business*, vol. 69, no. 3, pp. 279-311.

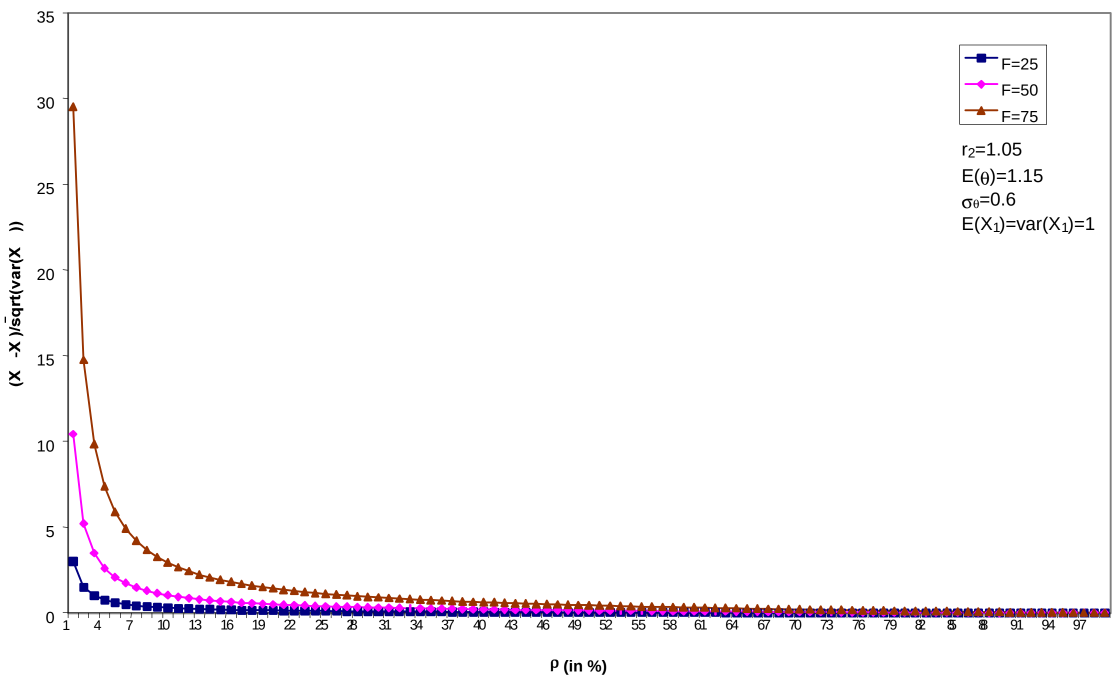
Titman, Sheridan and Roberto Wessels, 1988, “The Determinants of Capital Structure Choice”, *Journal of Finance*, vol. 43, no. 1, pp. 1-19, March.

Simulations

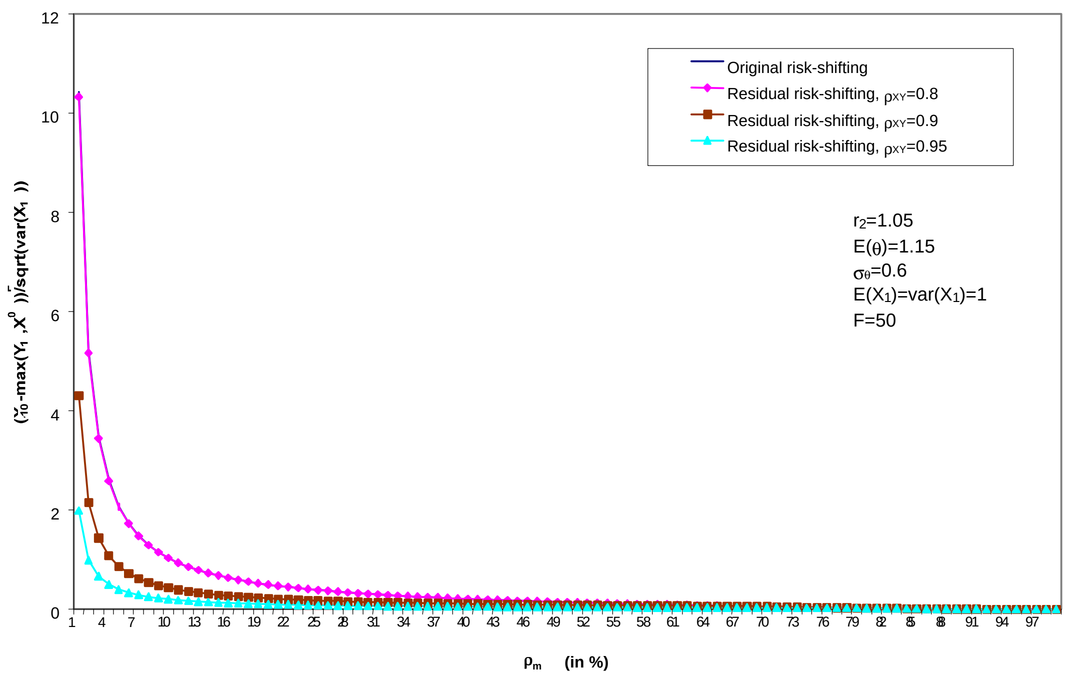
The investment cutoff as a function of the pattern of temporal resolution of uncertainty



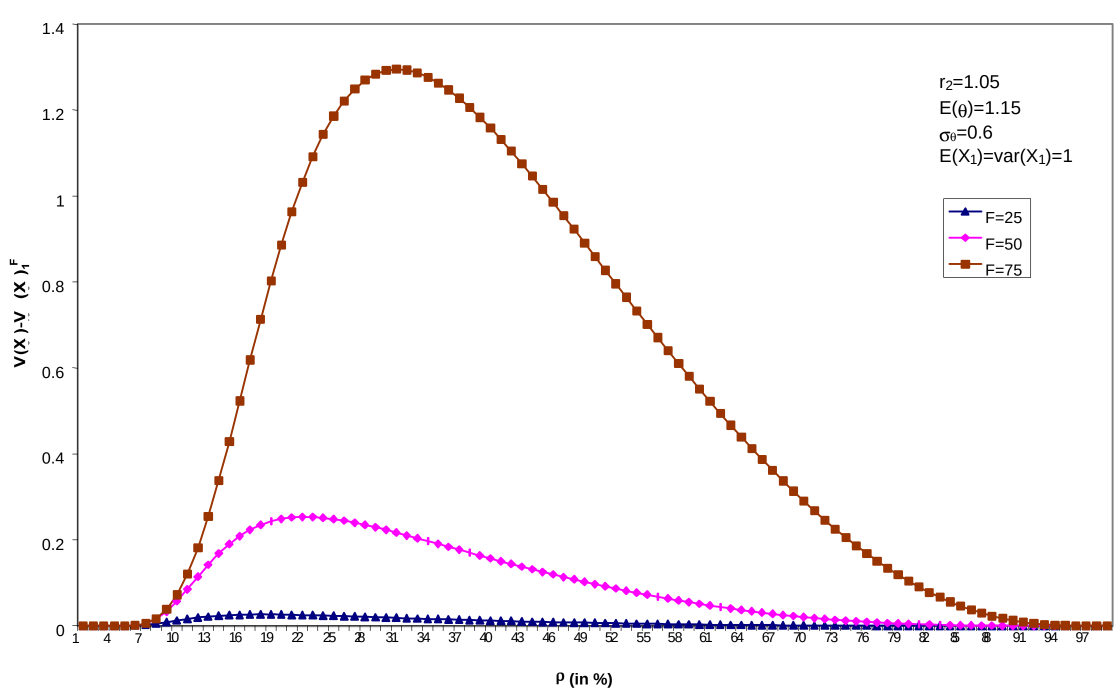
The extent of risk-shifting as a function of the pattern of temporal resolution of uncertainty



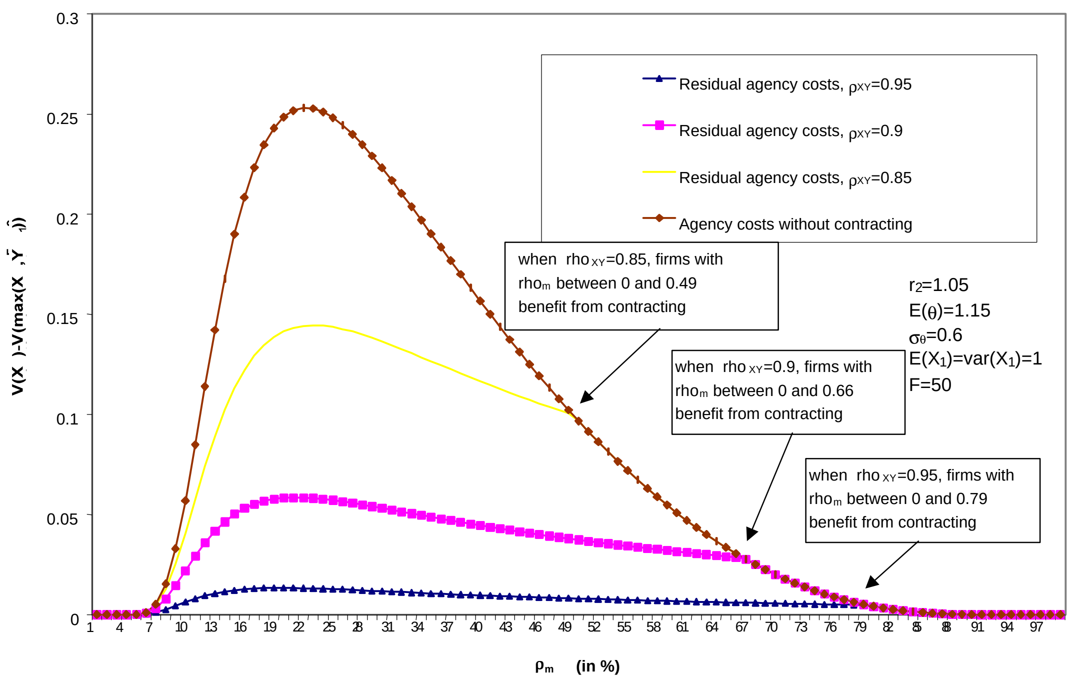
The extent of residual risk-shifting, once optimal contracting has been written, as a function of temporal resolution of uncertainty



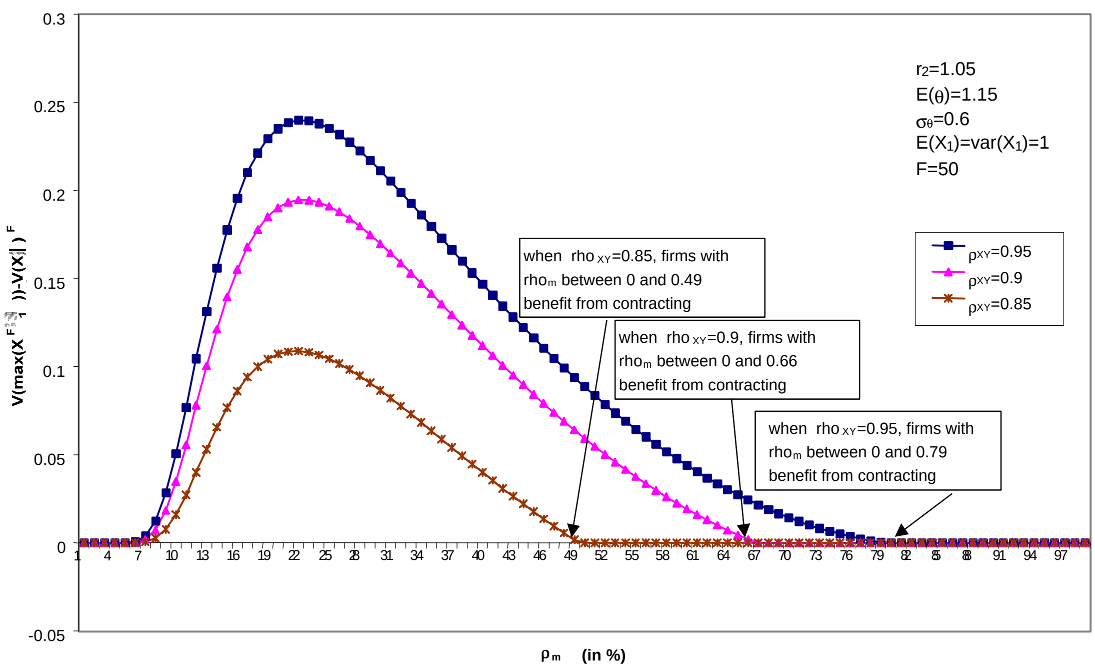
Agency costs as a function of the speed of resolution of uncertainty



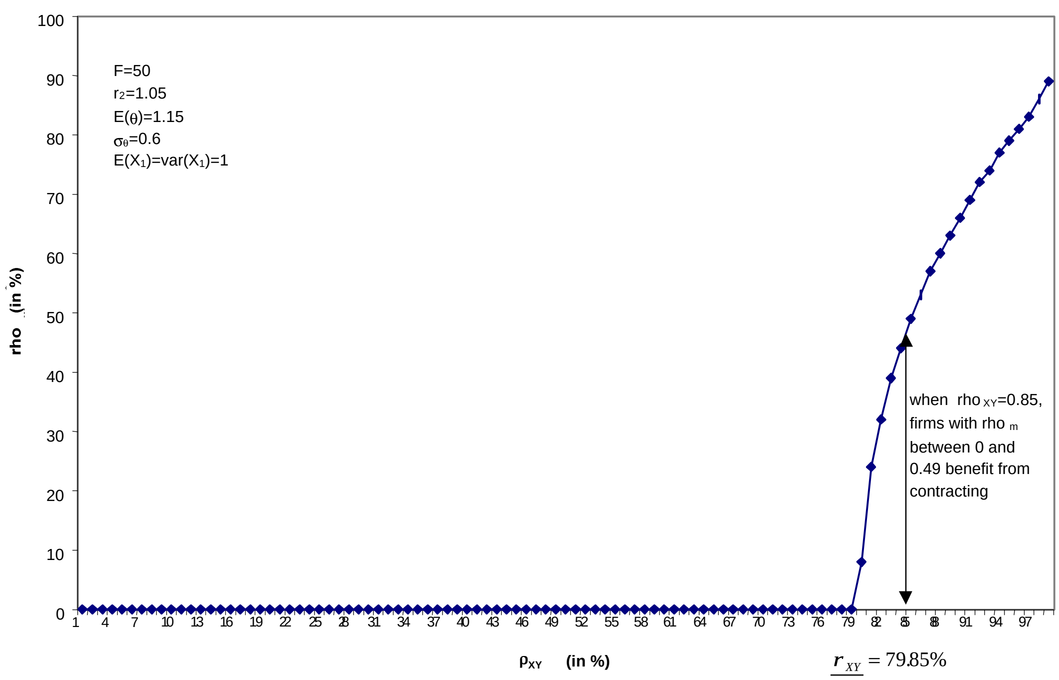
Residual agency costs after optimal contracting



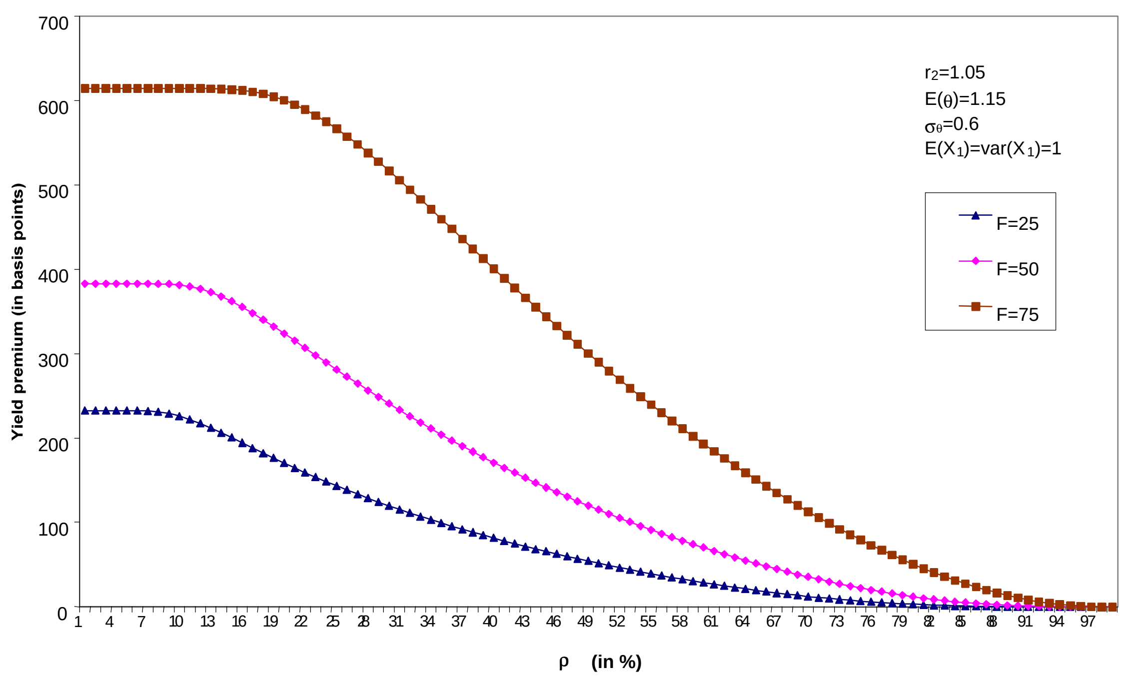
Improvement in agency costs as a function of the speed of temporal resolution of uncertainty



The maximum ρ_m^* for which contracting still presents benefits



Yield premium demanded on corporate bonds as a function of the pattern of temporal resolution of uncertainty



Yield premium demanded on bonds as a function of the pattern of temporal resolution of uncertainty